NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

FUSION OF IMAGES FROM DISSIMILAR SENSOR SYSTEMS

by

Khin Choong Chow

December 2004

Co Advisors: Monique P. Fargues
Alfred W. Cooper

Approved for public release; distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** December 2004	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE: Fusion of Images from Dissimilar Sensor Systems		**5. FUNDING NUMBERS**	
6. AUTHOR(S) Mr Khin Choong Chow, Republic of Singapore			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000		**8. PERFORMING ORGANIZATION REPORT NUMBER**	
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A		**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**	
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.		**12b. DISTRIBUTION CODE**	
13. ABSTRACT (maximum 200 words) Different sensors exploit different regions of the electromagnetic spectrum; therefore, a multi-sensor image fusion system can take full advantage of the complementary capabilities of individual sensors in the suit; to produce information that cannot be obtained by viewing the images separately. In this thesis, a framework for the multiresolution fusion of the night vision devices and thermal infrared imagery is presented. It encompasses a wavelet-based approach that supports both pixel-level and region-based fusion, and aims to maximize scene content by incorporating spectral information from both the source images. In pixel-level fusion, source images are decomposed into different scales, and salient directional features are extracted and selectively fused together by comparing the corresponding wavelet coefficients. To increase the degree of subject relevance in the fusion process, a region-based approach which uses a multiresolution segmentation algorithm to partition the image domain at different scales is proposed. The region's characteristics are then determined and used to guide the fusion process. The experimental results obtained demonstrate the feasibility of the approach. Potential applications of this development include improvements in night piloting (navigation and target discrimination), law enforcement etc.			
14. SUBJECT TERMS Image fusion, wavelet transform fusion, region-based fusion			**15. NUMBER OF PAGES** 93
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited.

FUSION OF IMAGES FROM DISSIMILAR SENSOR SYSTEMS

Khin Choong Chow
Civilian, Republic of Singapore
B.Eng, University College London, 1998

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN COMBAT SYSTEMS TECHNOLOGY

from the

NAVAL POSTGRADUATE SCHOOL
December 2004

Author: Khin Choong Chow

Approved by: Monique P. Fargues
Thesis Co–Advisor

Alfred W. Cooper
Thesis Co-Advisor

James H. Luscombe
Chairman, Department of Physics

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Different sensors exploit different regions of the electromagnetic spectrum; therefore, a multi-sensor image fusion system can take full advantage of the complementary capabilities of individual sensors in the suit; to produce information that cannot be obtained by viewing the images separately. In this thesis, a framework for the multiresolution fusion of the night vision devices and thermal infrared imagery is presented. It encompasses a wavelet-based approach that supports both pixel-level and region-based fusion, and aims to maximize scene content by incorporating spectral information from both the source images. In pixel-level fusion, source images are decomposed into different scales, and salient directional features are extracted and selectively fused together by comparing the corresponding wavelet coefficients. To increase the degree of subject relevance in the fusion process, a region-based approach which uses a multiresolution segmentation algorithm to partition the image domain at different scales is proposed. The region's characteristics are then determined and used to guide the fusion process. The experimental results obtained demonstrate the feasibility of the approach. Potential applications of this development include improvements in night piloting (navigation and target discrimination), law enforcement etc.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

The author would like to acknowledge and express his sincere gratitude to Professors Monique Fargues and Alfred Cooper for their wisdom, guidance and steadfast support of the completion of this thesis.

Support for this research was provided by Temasek Defence Systems Institute and National University of Singapore, under contract TDSI/02-005/1C – Multi-IR-Band Data Fusion for Target Recognition.

Special thanks to Dr. Dean Scribner, Naval Research Laboratory, Washington, DC for the provision of the NVD and Thermal Image file data set.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

The advent of night vision technology has increased the operational capabilities of modern armies by allowing soldiers to operate under the cover of darkness and poor visibility conditions [1]. In general, there are two classes of night vision technology: Night Vision Devices (NVD) and Thermal infrared (IR) systems. NVD enhance the very low levels of natural illumination, e.g. overcast star light, under which an unaided human eye would be essentially blind. IR sensors, in contrast, use heat emissions to identify objects that cannot otherwise be detected using available light sources. These systems support a wide range of military operations and have given the users a significant advantage over adversaries whose performance is degraded during night operation.

NVD and IR systems exploit different regions of the electromagnetic spectrum. Depending on the atmospheric and environmental conditions, one can offer better target information or situational awareness than the other. For example, NVD may have better image resolution but the contrast between heat-emitting objects and their surroundings is better in IR sensors, and therefore they offer a better dynamic range in detection. However, the information provided by each sensor is often complementary to the other; therefore limitations in each of the sensing modalities can sometimes be overcome by combining the input from multiple single-banded sources. This technique is known as multisensor fusion. It refers to the synergistic combination of different sources of sensory information into one representational format that is more suitable for human and machine perception or further image processing tasks. The information to be fused could come from multiple sensors monitoring over a common period of time or from a single sensor monitored over an extended period of time.

It has been shown that the joint use of imagery and spatial data from different imaging, mapping or other spatial sensors has the potential to provide significant performance improvements over single sensor detection, classification, and situation awareness. As a result, there has been a growing

interest in the use of multiple sensors to increase the capabilities of intelligent machines and systems, and multisensor fusion has become an area of intense research activity in the past few years.

This thesis seeks to improve the imagery produced by current night vision sensors by exploring different image processing techniques to combine the source images from NVD and IR sensors, and optimize the information content in the fused image. The image processing challenge is to develop an intuitively meaningful approach to extract the key features in each source image to facilitate the discrimination of objects from background and improve situational awareness.

This thesis is organized as follows: Chapter I covers the key motivations for undertaking this project. The next chapter describes the background to night vision and a review of the literature on image fusion. It also outlines the thesis objective. In Chapter III, wavelet transform theory, its application to image fusion and experimental results achieved are presented. Chapter IV introduces region-based fusion concepts and presents results demonstrating the robustness of the approach. Final remarks are provided in Chapter V.

II. BACKGROUND

A. NIGHT VISION

The human visual system is sensitive to radiation whose wavelength is in the 0.4 to 0.7 micrometer range of the electromagnetic spectrum. The visible radiation received by the human visual system depends on the amount of light present in the scene, or the luminance, and the amount of light reflected by object surfaces before reaching our eyes. When the scene illumination becomes low, our eyes lose color perception (due to the cone receptors) and objects appear in grayscale (scotopic vision).

Night vision technologies enable the exploitation of the night environment by processing the electromagnetic spectrum bands outside the human visual spectrum. The two bands exploited by NVD and IR imagers are the visible-near infrared band (wavelengths from 0.57 to 0.9 micrometer) and the thermal infrared band (wavelengths from 3 to 15 micrometer) respectively, as shown in Figure 1. The working principles for each sensor system are summarized in the following two sub-sections.

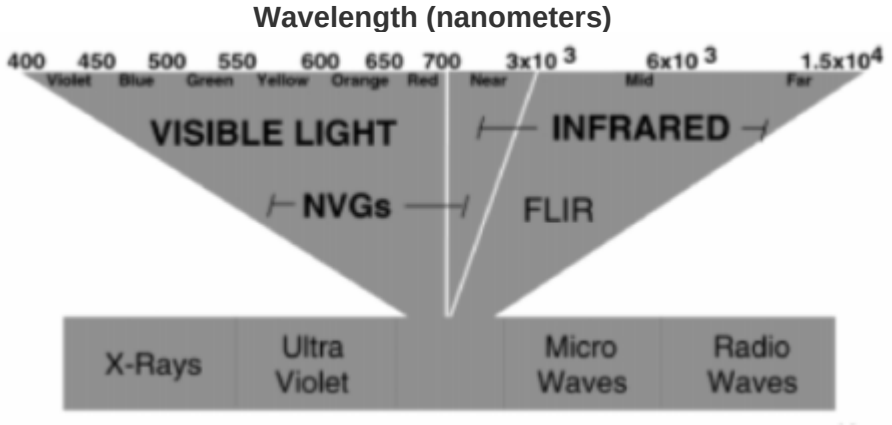


Figure 1. Spectral response of the eye, NVD and thermal IR sensors (From Ref. [2]).

1. Night Vision Devices

NVDs are passive devices that operate in the visible and near-infrared regions of the electromagnetic spectrum (Figure 1). Much like the human visual system, they depend almost entirely on the reflected energy from the scene illumination. Including an image intensifier in the optical system amplifies the very low radiance of natural light that is reflected by the scene (target and background). Image intensifiers are classified in three categories: first, second and third generation, each with different performance characteristics. A typical night vision goggle (Generation II or III) assembly consists of an objective lens, photocathode, microchannel plate, phosphor screen and combiner eyepiece assembly (Figure 2).

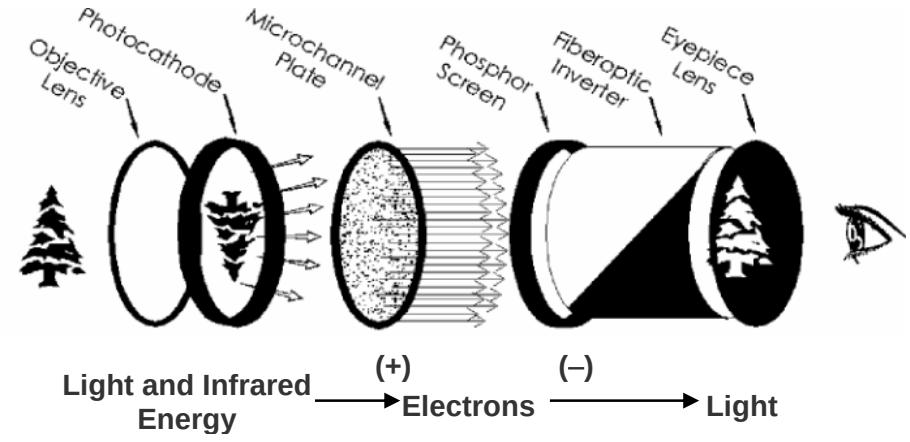


Figure 2. Night vision device with microchannel plate to collimate electron flow and increase the light-amplification gain (From Ref. [2]).

Radiant or reflected optical energy received at this device is focused by the objective lens onto the photocathode. The photocathode, which is responsive to both visible and near-IR radiation, converts the incident photons into photoelectrons. The released electrons are then accelerated by an applied electric field through a microchannel plate. Successive secondary electron emission occurs in the pores of the microchannel plate leading to multiplication by a factor of up to four orders of magnitude. These electrons are further accelerated to strike a phosphor screen which in turn coverts the high energy

electrons back to light (photons), which corresponds to the distribution of the input image radiation but with a flux amplified many times [3].

Visible and near-infrared night-time imagery is currently provided by the third generation of image intensifier tubes. The variants of the Gen III NVGs currently used have a gain in the order of magnitude of 30,000 to 70,000.

2. Thermal Infrared Devices

Thermal infrared devices detect invisible self-radiating and reflected infrared (IR) radiation from objects in the scene and convert this energy into a visible image. The infrared range covers all electromagnetic radiation from 0.7 to 20 micrometer. However, only certain “atmospheric windows” exist (Figure 3). This is due to the absorption of the radiation by different gases and water vapour in the atmosphere. Therefore, the two bands that are generally employed by forward looking infrared sensors (FLIR) are the medium wavelength IR (MWIR - 3 to 5 micrometer) and long wavelength IR (LWIR - 8 to 12 micrometer).

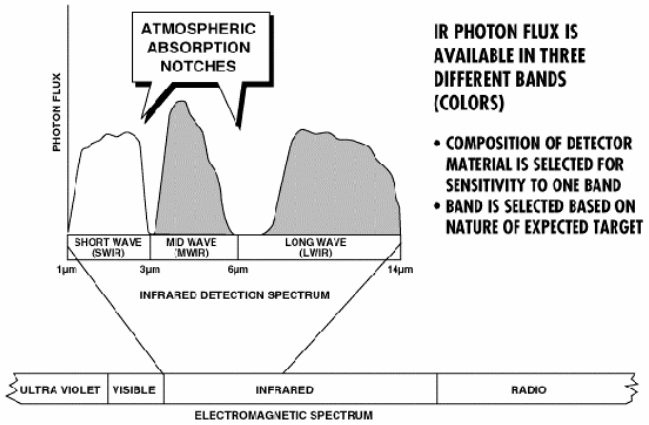


Figure 3. IR spectral bands and atmospheric transmittance as a function of wavelength. The “atmospheric windows” are the gaps between the absorption regions due to different gas and water vapour molecules in the atmosphere (From Ref. [4]).

All objects are composed of continually vibrating atoms. The vibration of all charged particles, including the electronic structure of these atoms generates electromagnetic waves. The electromagnetic radiation is emitted with a wavelength distribution at a rate that depends upon the temperature of the object and its spectral emissivity. Emissivity compares the ability of a material to emit IR

energy to that of a blackbody at the same temperature. A “Blackbody” is defined as the perfect absorber of thermal energy and therefore also a perfect emitter, with an efficiency of unity. It is a function of both the type and surface finish of the material. Figure 4 shows how the energy emitted increases with temperature [5].

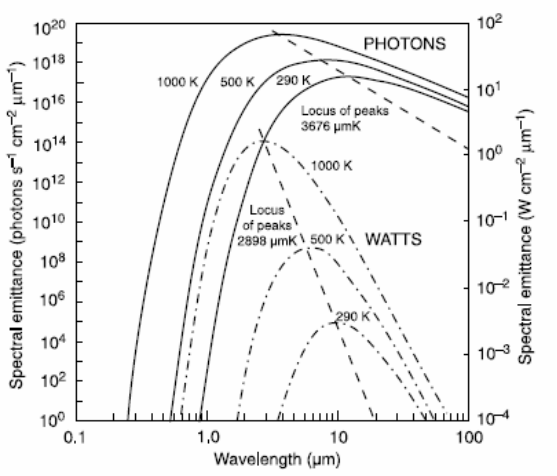


Figure 4. Planck’s law for spectral emittance (From Ref. [5]).

The thermal signature of an object is determined by the thermal flux self-generated or reflected from other heat sources. Humans, animals and objects in nature frequently have a high emissivity and therefore, a majority of their signature is from self-emission, which at normal temperature tends to peak in the LWIR band. Conversely, objects with low emissivity have a corresponding high reflectivity and therefore, reflect the thermal energy of their surroundings, e.g. solar scattered radiation, which is significant only by day and has a maximum emission in the MWIR band. A body with high reflectivity in one wave band may have high emissivity in another.

Modern infrared detectors generally fall into two categories, Photon and Thermal detectors. In photon detectors, the radiation is absorbed within the material to produce electrons, which can be detected as voltage or current. They exhibit both high sensitivity and a very fast response, and the response per unit incident radiant power is wavelength dependent. However, photon detectors for the thermal IR are generally required to be cooled to very low temperatures,

typically 77K during operation, making them bulky and expensive. They are usually used in high performance systems.

Thermal detectors work on the principle that the incident radiation heats up the material of the detector and causes a change in some physical property, e.g. resistance, which can be detected as an electrical output. They are generally wavelength independent and characterized by modest sensitivity and slow response.

3 Comparison Between NVD and Thermal IR Imagery

Figure 5 and Figure 6 show the image of the same scene captured by a NVD and a thermal IR camera. In the NVD image, the low night sky lighting reflected in the environment is amplified by the image intensifier to give a low contrast image with limited dynamic contrast range. As a result, the night sky and the ground terrain, including the track in the lower portion of the image, are captured with limited details. Despite the limited contrast range, the treeline can be differentiated clearly as the night sky is better illuminated. Two bright artificial self-emitting light sources are also captured in the image.

The thermal IR image, on the other hand, reflects greater details or “texture” in the foreground. This is due to greater contrast in emissivity between the track and its adjacent terrain. However, the similarity in temperature between the night sky and the distant treeline resulted in an almost uniform continuation between the two regions. This could be partly attributed to lower resolution of the thermal IR camera, which fails to capture the minor temperature variation in the far field. Lastly, the two artificial light sources emit radiation in the shorter wavelengths which are beyond the bandwidth of the thermal IR camera. Therefore, they are not captured in the thermal IR image.

The two images presented capture the different details in the scene as they operate in different regions of the electromagnetic spectrum. The complementary set of images suggests the feasibility of combining the source images into a fused image that aims to increase the scene content.

Figure 5. Image captured by NVD (From Naval Research Laboratory).

Figure 6. Image of the same scene captured by a thermal IR camera (From Naval Research Laboratory).

B. REVIEW OF THE LITERATURE IN SENSOR FUSION

The objective of image fusion is to generate a hybrid high-resolution multi-spectral representation that attempts to preserve the radiometric characteristics of the original multi-spectral data. Various fusion approaches have been proposed for the merging of multi-spectral and high spatial resolution data, including “statistical and numerical” and “multiresolution analysis” methods.

Image fusion by the statistical and numerical approach utilizes methods such as Principal Component Analysis (PCA) and Principal Component Substitution to extract key information from the disparate sensor inputs. This forms the basis for the fusion process. In the Naval Research Laboratory’s color fusion algorithm [6], a red-green color opponency was used to display a dual band infrared image (Figure 7). L_1 and L_2 represent the pixel intensities in LWIR and MWIR sensors respectively. They are statistically decomposed using PCA into orthogonal components L_1’ and L_2’, which correspond to the brightness and chromatic axis respectively. The distribution along the brightness axis represents a high correlation in intensity distribution between the pixels while the orthogonal component L_2’ maps to the uncorrelated pixel intensities.

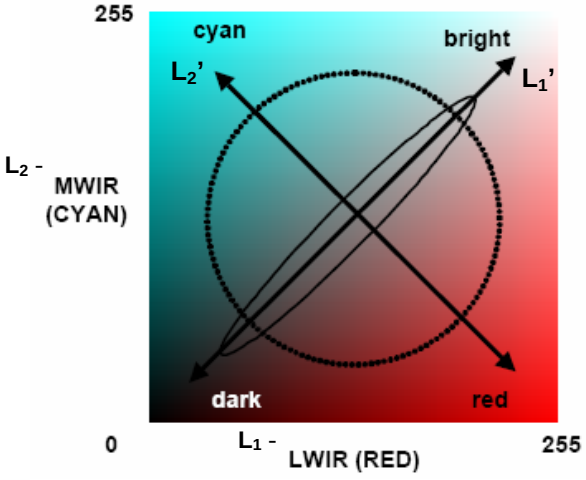


Figure 7. Principal component direction (brightness) and its orthogonal principal component (chromaticity plane) (From Ref. [6]).

In the fused image, each pixel is assigned a chrominant value (red-cyan) and brightness value (black-white), depending to the location of the input pixel intensity pair relative to two principal components. Therefore, features that are

present in both the sensors are represented by grayscale intensity as they have a corresponding pixel intensity pair that is close to the brightness axis. Conversely, features unique to each sensor have pixel intensity pairs that are far away from the brightness axis and are distributed along the principal component L_2'. They are represented in a red-cyan combination.

Another approach for fusing low-light visible and uncooled thermal infrared imager data is proposed in [7]. In the paper, Therrien et al. describe an enhanced Peli-Lim algorithm to perform adaptive modification of the local contrast and local luminance mean, which is accomplished by separating the source images into spatial high-pass (local contrast) and low-pass components (local luminance mean). The high-pass components are enhanced by multiplying them by a gain factor that depends on the local luminance mean while low-pass components are passed through a nonlinear luminance transformation to reduce their dynamic range. The local energies of the high-pass components from the input sensors are then computed. The images are fused using a weighted combination of the source images based on a normalized difference of local energies. The block diagram of the enhanced Peli-Lim algorithm is shown in Figure 8.

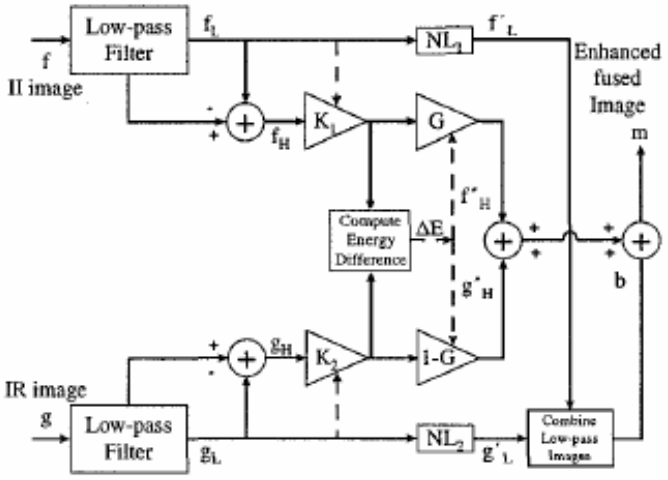


Figure 8. Block diagram of the enhanced Peli-Lim algorithm (From Ref. [7]).

In [8], Qu et al. noted that the spectral characteristics of the source images are not well preserved in color transformation, statistical, and numerical methods as they tend to alter the fused image features. Therefore, these methods have been replaced by fusion schemes based on multiresolution decomposition. Another motivation for pursuing the multiresolution approach lies in the fact that real-world scenario contains objects or features of different sizes. As a result, performing image analysis at a single scale tends to ignore the features that are present at other scales and this may result in the loss of spectral information in the fused image. The solution is to adopt a multiresolution approach that analyzes the image at different scales.

One of the earliest multiresolution approaches is the pyramid decomposition scheme, first proposed by Burt [9,10]. In a Gaussian pyramid, the original image is repeatedly filtered and sub-sampled to generate the sequence of reduced resolution sub-images. This approach is equivalent to convolving the original image with a set of Gaussian-like weighting functions, followed by sub-sampling. In a Laplacian pyramid, the sub-image at each level of the pyramid is given by the difference between successive levels of the Gaussian pyramid. In image fusion, a pyramid transform is constructed for each input image. The pyramid image is then combined using some selection rule to form a composite image pyramid. Finally, the fused image is recovered by taking an inverse pyramid transform of the composite pyramid.

In [11], Li et al. noted that pyramid-based techniques result in redundancy between different resolutions and merged images contain blocking effects in the regions where the input data from different sensors are significantly different. Therefore, multiresolution wavelet-based methods have been proposed. Wavelets are functions defined over a finite interval. The basic idea is to represent an arbitrary function as a linear combination of a set of such wavelets or functions. Over the last few years, the wavelet transform has been widely used in image fusion applications to fuse multimodal sensor data into a composite representation. In many applications, the wavelet-based approach works well in preserving the spectral information of the source images.

C. OBJECTIVE

The information provided by different sensors is often complementary, therefore improvements are possible with the enhancement and subsequent fusion of the images captured into a single representation. Among the different fusion schemes, the multiresolution approach based on the wavelet transform offers one of the most promising solutions to effectively extract and combine the salient features in the source images. By analyzing and fusing the source images at different scales, the wavelet-based technique provides a more reliable means to preserve the spectral information of the multispectral images.

Therefore, this thesis seeks to implement a wavelet-based image fusion algorithm to fuse images received from dissimilar image sensors, in particular, complementary images from thermal and night vision sensor systems. In the wavelet domain, many image processing techniques, e.g., denoising, contrast enhancement, segmentation, texture analysis and compression can be easily performed. In addition, this thesis also explores other pre-processing techniques to improve the fusion results.

III. WAVELET TRANSFORM FUSION

A. OVERVIEW

Image fusion can be defined as the process of combining multimodal source images into a single representation, emphasizing the most salient features of the surrounding environment. According to [12], an image fusion algorithm should preserve as closely as possible all relevant information contained in the source images and not introduce any artifacts or inconsistencies that could interfere with interpretation. In the fused image, the irrelevant features or noise should also be suppressed to a maximum extent.

The actual fusion process can take place at different levels of information representation. These approaches fall into three basic categories, i.e. pixel, feature and decision level fusion [13]. At the lowest processing pixel level, the sets of pixels in the source images are merged pixel to pixel according to a defined decision rule to form the corresponding pixel in the fused image. Fusion at this level requires accurate spatial registration of the images from different sensors prior to applying the fusion operator. In feature level fusion, the relevant features are first abstracted from the data and then fused to form the fused feature set. The features can be extracted using segmentation procedures and differentiated by characteristics such as size, shape, contrast and texture. As the fusion is based on identified features in the sources, the resulting probability of detecting useful features in the fused image increases. At the decision level, decisions/detections based on the outputs from the individual sensors are fused together and used to reinforce common interpretation or resolve any differences.

Among these three fusion methods, pixel level fusion is the most mature, as it has the advantage of directly using the source images that contain the original information. In addition, the algorithms used are also typically more time efficient. They range from the simple image averaging type to the complex PCA, pyramid-based image fusion and wavelet transform fusion.

The following section will include: 1) a brief overview of wavelet analysis, the discrete wavelet transform and its implementation to serve as a prelude to the development of the fusion technique; 2) a description of the image analysis using discrete wavelet transform, and 3) the theory and experimental results of wavelet transform image fusion using different fusion rules.

B WAVELET TRANSFORM

The fundamental idea behind the wavelet transform is to analyze a signal at different scales or resolutions. The wavelet transform can be interpreted in the Fourier domain as set of band-pass filters and the signal is examined in both the space and frequency domains. Its transform allows a signal *f(t)* to be projected onto different wavelets or basis functions instead of the sin and cosine basis functions that are used in Fourier transform. These basis functions are obtained from a single prototype wavelet called the mother wavelet by dilations and translations. In the wavelet domain, the larger wavelets give the approximate signal representation while the smaller wavelets zoom in to the details or minor variations in the signal.

While sinusoids are useful in analyzing periodic and time-invariant phenomena, wavelets are well suited for the analysis of transient, time-varying signals. The great interest in the use of wavelets for signal and image analysis lies in the ability to efficiently represent functions with localized features. Compared to pyramid transforms, discrete wavelet transform is also more compact and offers directional information [12]. In image analysis, the 1-dimensional wavelet transform is extended to the 2-dimensional wavelet transform to perform spatial-frequency decomposition of the source image.

1. Continuous Wavelet Transform

The basic idea of wavelet transform is to represent any arbitrary function as a decomposition in terms of the basis functions. For a one-dimensional signal *f(t)*, the continuous wavelet transform is defined using the relation [14]

$$W_{\psi}(f)(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \psi^{*} \left(\frac{t-b}{a} \right) dt, \tag{3.1}$$

where $\psi(t)$ is the mother wavelet, *a* is the scaling factor, and *b* is the shifting factor. The wavelet coefficient $W_{\psi}(f)(a,b)$ provides the information on the signal at each location *b* and for the scale *a*. Reconstruction can be obtained from the wavelet coefficients by using the inverse wavelet transform:

$$f(t) = \frac{1}{C_{\psi}} \int_{-\infty}^{+\infty} db \int_{-\infty}^{+\infty} \frac{1}{\sqrt{a}} W_{\psi}(f)(a,b) \psi \left(\frac{t-b}{a} \right) db, \tag{3.2}$$

where C_{ψ} is a factor that depends on the choice of wavelet and is given by:

$$C_{\psi} = \int_{-\infty}^{+\infty} \frac{\left| \Psi(w)^2 \right|}{w} dw < \infty, \tag{3.3}$$

and $\Psi(w)$ is the Fourier transform of $\psi(t)$.

2. Discrete Wavelet Transform

Continuous wavelet transform places redundant information on the time-frequency plane and is computationally expensive. Therefore, the discrete wavelet transform (DWT) was developed to analyze a signal using a subset of scales and positions.

According to [14] and [15], the wavelet decomposition of a discrete signal *f(t)* is given as:

$$f(t) = \sum_{m} \sum_{n} c_{m,n} \psi_{mn}(t), \tag{3.4}$$

where *m* and *n* are integers and $\psi_{mn}(t)$ is a wavelet basis function. The two-parameter DWT coefficient c_{mn} is given by:

$$c_{mn} = \int f(t) \psi_{mn}(t) dt. \tag{3.5}$$

The wavelet basis function, $\psi_{jk}(t)$ relates to the mother wavelet $\psi(t)$ by the following relation:

$$\psi_{mn}(t) = 2^{m/2}\psi(2^m t - n), \tag{3.6}$$

where *n* is the translation and *m* the dilation parameter. Equation (3.6) shows that the wavelet basis functions are formed by translating and scaling the mother wavelet. An additional set of coefficients, a_{mn} is used to describe the trend or approximation of the function *f(t)* at resolution 2^m during a recursive wavelet transform. The difference between one approximation and the other at the next level is known as "detail", and is given by the wavelet coefficient c_{mn}.

Wavelet families have different properties and differ in terms of the basis functions compactness, spatial localization, and smoothness; hence they are suitable for different applications. The Haar, Daubechies, Symlets and Coiflets wavelets are examples of orthogonal wavelet families that remove the correlation in the signal between different subspaces, and hence avoid redundancy in the decomposed signal representation between different resolutions. Figure 9 shows the above four wavelet families.

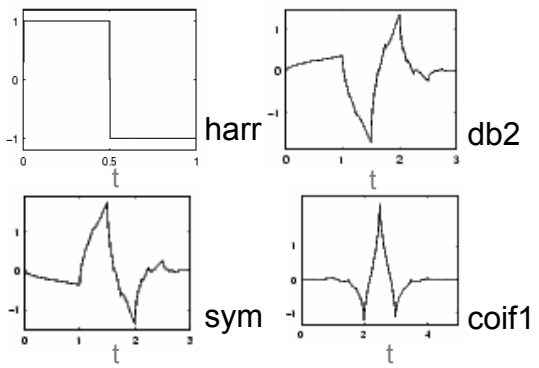


Figure 9. Wavelet families – Haar, Daubechies-2, Symmlet and Coiflet (From Ref. [16]).

The Haar wavelet transform is the simplest transform to implement. It allows quick visual inspection of the wavelet levels. However, a major disadvantage is its discontinuity, which makes it difficult to represent a continuous signal.

Ingrid Daubechies invented the first continuous orthogonal compact support wavelet, the Daubechies wavelet. It is suitable for continuous transform and has been widely used in signal and image analysis applications. The Symmlet and Coiflet have near symmetry properties, which allows the corresponding wavelet transform to be implemented using minor boundary conditions that can reduce boundary artifacts [16].

In this thesis, one of the most commonly applied and proven wavelet families, Daubechies wavelets, will be used to develop the framework for the wavelet-based image fusion scheme. Once the framework is developed, other wavelet families, e.g. Symmlet and Coiflet wavelets may be explored to determine the optimal wavelet selection for the fusion of NVD and thermal IR images.

3. Image Analysis Using Discrete Wavelet Transform

In general, an image comprises features or objects at different scales. Therefore, multiresolution techniques were developed to extract scale-specific information from the image, in particular, coarse scale information in high levels and fine scale information in low decomposition levels. The DWT provides a framework for such multiresolution image analysis. The 1-dimensional DWT can be extended to a 2-dimensional DWT to perform spatial-frequency decomposition on a source image into a multiresolution pyramid of new images.

In [17], Mallat introduced a fast discrete 2-dimensional wavelet transform algorithm that is based on the use of multiresolution approach for image analysis. The transform can be implemented recursively using a set of low-pass finite impulse response (FIR) filters h_n and related high-pass FIR filters g_n to derive the approximate (a_{mn}) and details (c_{mn}) coefficients, respectively. The 2-dimensional data is separately filtered and downsampled in the horizontal and vertical direction to produce four sub-bands at each scale, as illustrated in Figure 10.

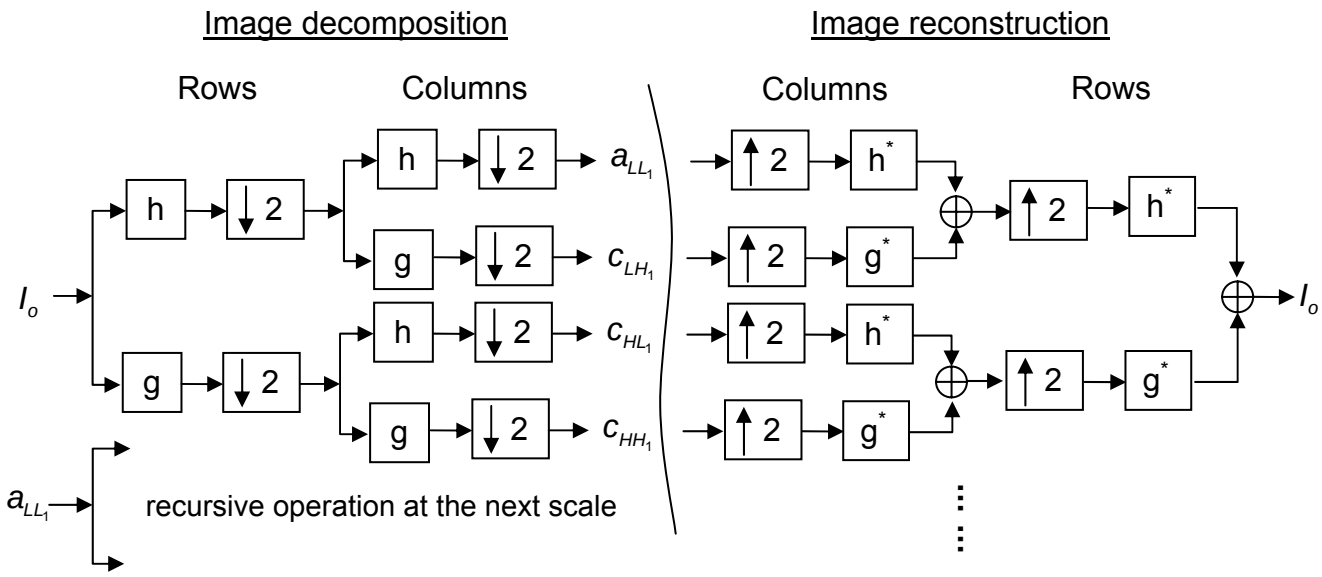


Figure 10. 2-dimensional wavelet transform using filter operations. The input I_o is decomposed into four sub-images corresponding to the approximate image a_{LL_1} and detail images c_{LH_1}, c_{HL_1} and c_{HH_1}. Subsequent reconstruction produces the input image.

Therefore, given a grayscale input image I_o, the 2-D wavelet decomposition gives

$$I_o = a_{LL_1} + c_{LH_1} + c_{HL_1} + c_{HH_1}, \tag{3.7}$$

where the sub-image approximation a_{LL_1} is the base low frequency image. It represents the averaged, lower resolution version of the image I_o. The detail sub-images correspond to the high frequency parts or features of the image. They contain information about I_o not present in the simplified component a_{LL_1}. c_{LH_1} tends to emphasize the horizontal edges and is referred to as the first horizontal fluctuation while c_{HL_1} is known as the vertical fluctuation as it emphasizes the vertical edges. The last detail, c_{HH_1} represents the first diagonal fluctuation and tends to emphasize the image diagonal features.

The first approximate sub-image is then decomposed to the next level:

$$a_{LL_1} = a_{LL_2} + c_{LH_2} + c_{HL_2} + c_{HH_2}. \tag{3.8}$$

Recursively, by taking successive approximations of the original image at increasing scales in the wavelet transform, an image pyramid is formed. At the n^{th} level, it will comprise *3n+1* sub-image sequences. Each input image can be decomposed up to the maximum decomposition level, which is $\log_2 N - 1$ (*M* by *N* = size of the image, $N \leq M$). Figure 11 shows the image sub-bands in the decomposition process. Note that by applying inverse wavelet transform, the n^{th} level approximate image a_{LL_n} can be perfectly reconstructed from the $(n+1)^{th}$ level coefficients, $a_{LL_{n+1}}, c_{LH_{n+1}}, c_{HL_{n+1}}$ and $c_{HH_{n+1}}$ by means of backward recursion.

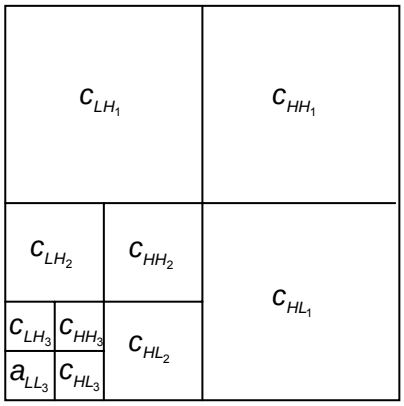


Figure 11. Image sub-bands.

Figure 12 to Figure 15 illustrate the concept of multiresolution wavelet decomposition. Downsampled representations consisting of one approximate and three detail sub-images are generated at every level of the decomposition. The approximate sub-images A1, A2 and A3 represent a lower resolution approximation of the original image and they retain some of its properties such as the mean intensity or texture information. In the detail sub-images, the horizontal, vertical and diagonal fluctuations are picked up by the respective detail coefficients at each scale. For example, horizontal roof edge and steps are captured in the horizontal detail sub-images while vertical pillars and edges of the wall are reflected in the vertical detail sub-images.

The illustrations show that the finer details are captured in the lower levels of decomposition while the coarser scale information is presented in the higher levels of decomposition. It also demonstrates that the multiresolution wavelet transform is able to identify the salient directional features in an input image. This highlights the feasibility of fusing images from different sensors by combining the key features identified at each scale. It is further motivated by the fact that the human visual system is primarily sensitive to local contrast changes such as edges or corners and the improved scene content will aid situation awareness and scene recognition.

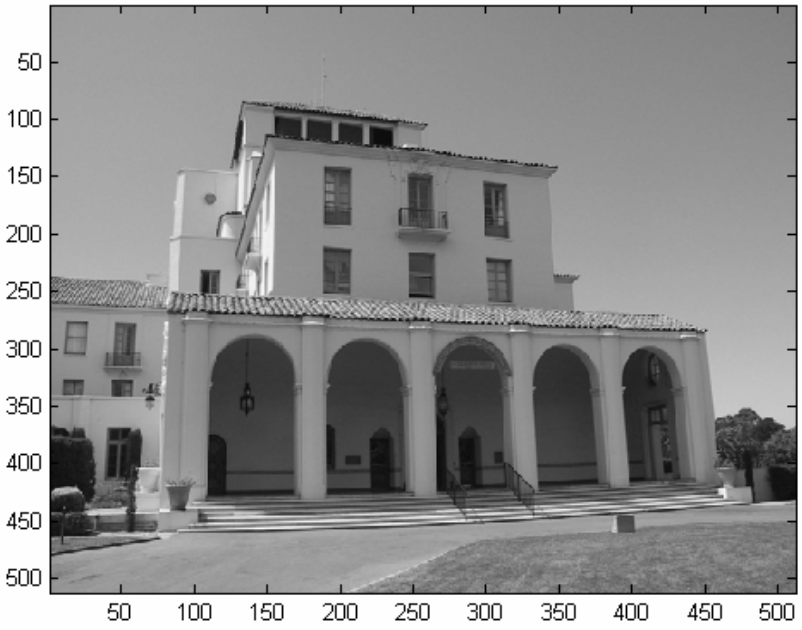


Figure 12. Original Image – Herrmann Hall, NPS.

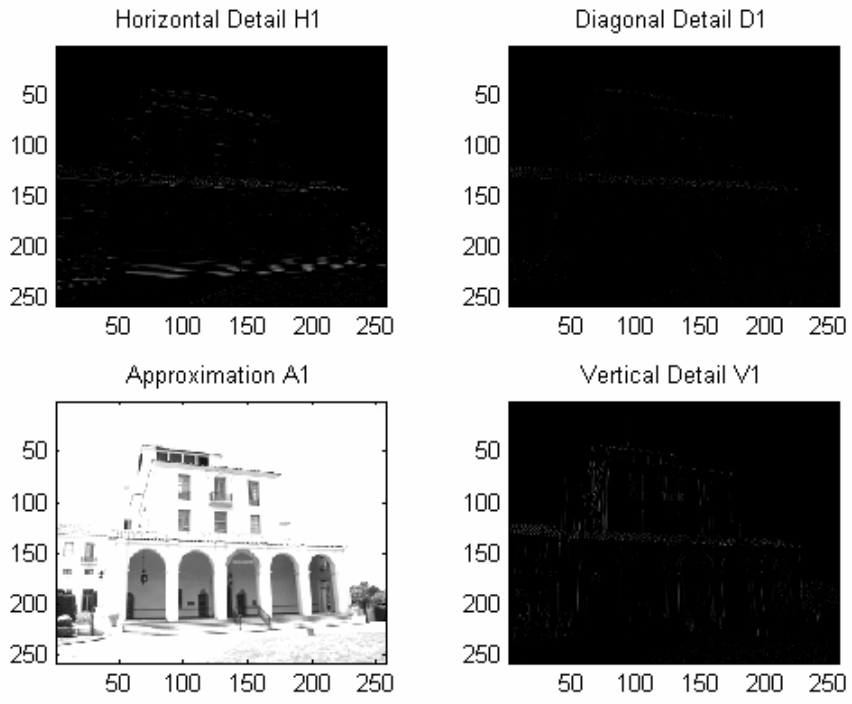


Figure 13. Wavelet decomposition at level 1. The approximate sub-image is a coarse representation of the original image and the horizontal, diagonal and vertical variations are captured in the detail sub-images.

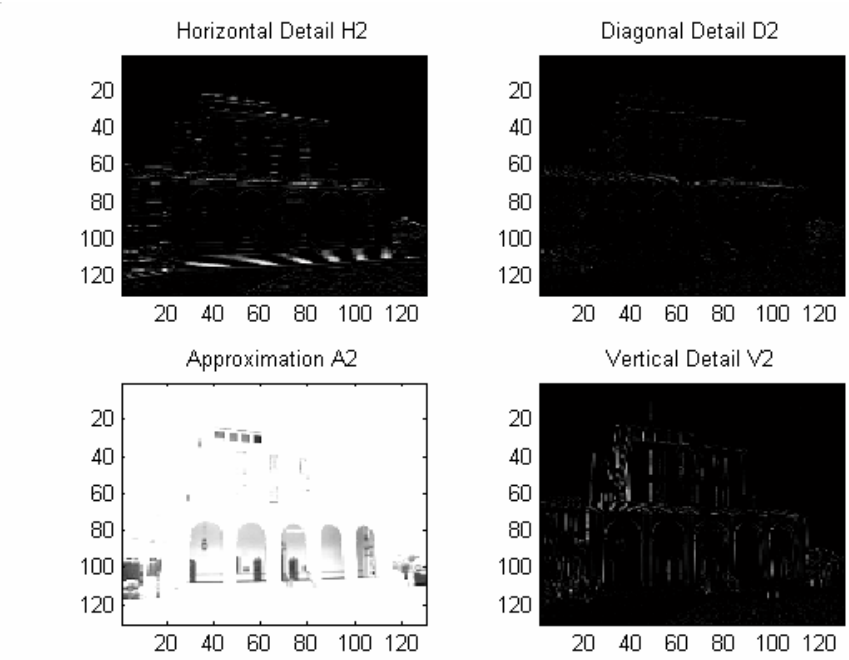


Figure 14. Wavelet decomposition at level 2. The lower resolution sub-images A2, H2, D2 and V2 are derived from the level 1 approximate sub-image A1. Notice how they capture the salient features in the original image at a coarser scale.

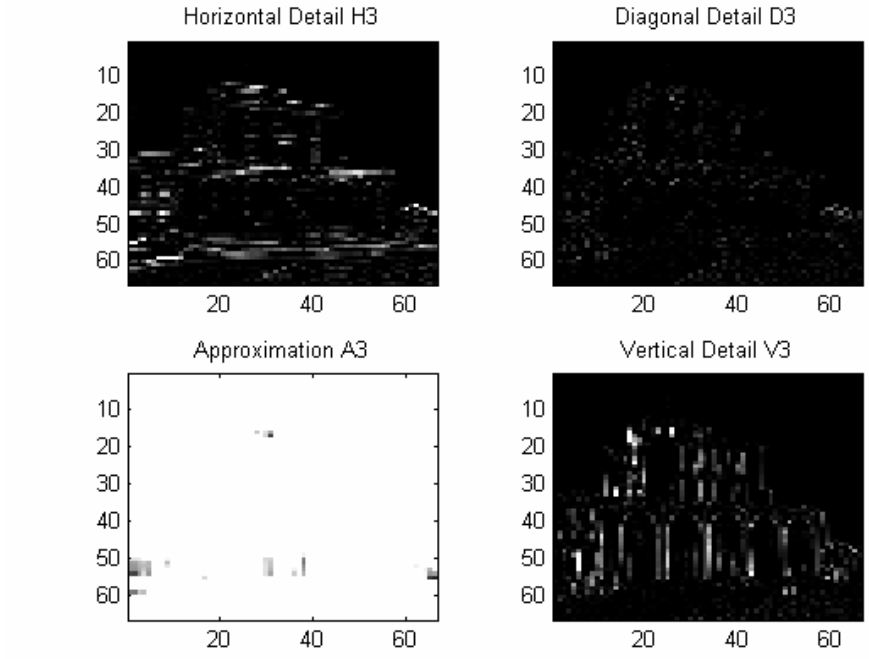


Figure 15. Wavelet decomposition at level 3. The lowest resolution sub-images are presented at this level.

C. WAVELET TRANSFORM FUSION

The principle of image fusion using wavelet-based decomposition is to selectively merge the decomposed "approximation" and "details" coefficients of the original images. An inverse transform performed on the fused coefficient representation will give the fused composite image. There exist many variations in the approach for multiresolution fusion [9,10,11,18].

The general framework for the multiresolution wavelet transform fusion scheme is presented in Figure 16. Application of this framework to a set of registered source images will produce the fused output. At each level of decomposition, a decision that is governed by a set of fusion rules is made to decide how the multiscale representations should be used to construct the fused wavelet coefficient map.

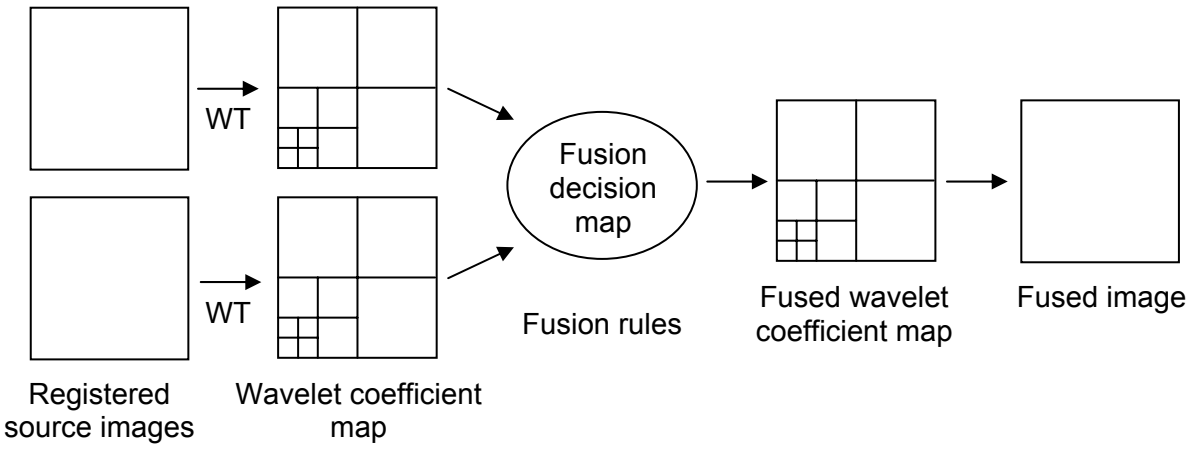


Figure 16. General framework for image fusion using multiresolution wavelet transform. Registered source images are decomposed, fused according to the fusion rule and reconstructed to produce the fused image (After Ref. [11]).

Pixel-based image fusion requires the source images to be aligned on a pixel-by-pixel basis. The techniques for image registration are widely researched and discussed in the literature and therefore will not be covered here. It is assumed that the images to be combined are registered.

1. Fusion Rules

Since the salient features are captured by the detail wavelet coefficients, the key of successful fusion lies in defining an appropriate feature selection fusion rule to select and construct the fused detail wavelet coefficient maps at each scale. A more detailed illustration of the framework for the formation of the fusion decision map is shown in Figure 17.

The framework uses an activity and matching measure to define the fusion rules, which will then be used to generate the fusion decision map. The output of the decision map will govern the actual combination of the coefficients from the wavelet decompositions of the source images.

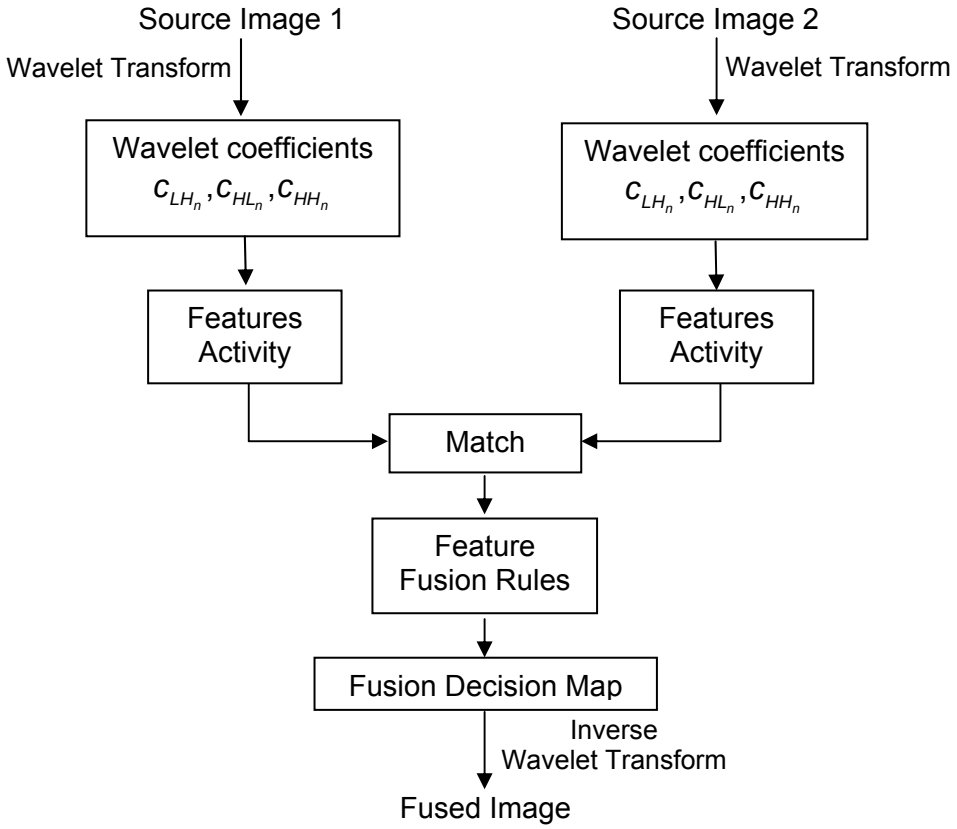


Figure 17. Framework for the Formation of the Fusion decision map.

As the approximate sub-image represents the coarse approximation to the original image, the most common approach used to derive the fused approximate wavelet coefficient map is by taking the average of the source images' approximate coefficients at each level.

The activity-level measurement reflects the salience of a particular pixel in the image. It is high if the pixel represents important information in a scene; conversely, it is low if the pixel represents some unimportant information. Two methods are used to determine this activity level. In general, a pixel is expected to be important if it is relatively prominent in the image. Therefore, in the simpler case, the larger absolute value of the details wavelet coefficient can be used as a generic measure of its salience. It is given by

$$a_{A_n}(j,k)=\left|c_{A_n}(j,k)\right| \text{ and } a_{B_n}(j,k)=\left|c_{B_n}(j,k)\right|, \tag{3.9}$$

where $c_{A_n}(j,k)$ and $c_{B_n}(j,k)$ represent the n^{th} level wavelet coefficients at location (*j,k*) of input image A and B, respectively. This is known as the pixel-based method [18].

The second method considers a neighborhood pattern around the sampled pixel. It takes into account that the surrounding pixels would be highly correlated to the sampled pixel if it represents a salient feature. Typically, a 3 by 3 or 5 by 5 window centered at the sampled pixel is used [19]. This method is known as the window-based activity measure and can be implemented as:

$$a_{W(A,B_n)}(j,k)= \sum_{s\in S,t\in T} c_{A,B_n}(j+s,k+t), \tag{3.10}$$

where *S* and *T* are sets of horizontal and vertical indexes that describe the current window. It measures the activity associated with the n^{th} level pixel centered in the window at location (*j,k*). Increasing the size of the neighborhood will add robustness to the fusion system as it will reduce the contribution of localized noise at higher computational cost. At lower resolutions of decomposition, the window may also exceed the size of the local features. Figure 18 illustrates the differences between pixel and window-based fusion rules.

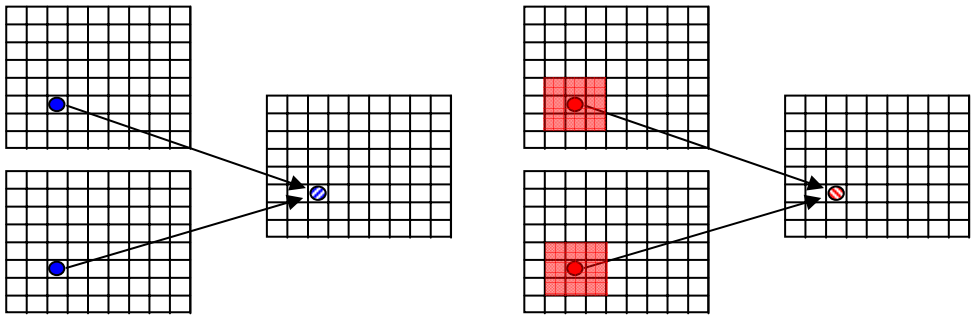


Figure 18. Comparison between pixel and window based fusion rules.

The matching measure is used to determine the degree of resemblance between corresponding pixels in the source images and this information will be used to determine the mode of combination at each pixel location. It is given by the correlation between the corresponding pixels at location (*j,k*) for the n^{th} level coefficients:

$$m_{F_n}(j,k)= \frac{2c_{A_n}(j,k)c_{B_n}(j,k)}{\left|c_{A_n}(j,k)\right|^2+\left|c_{B_n}(j,k)\right|^2}. \tag{3.11}$$

Several different DWT-based fusion rule schemes have been proposed in [18,10,19,etc.]. In this thesis, three fusion rules are implemented.

Fusion Rule 1 – Selection of the dominant mode

Using the parameters defined above, the simplest fusion rule is to select the coefficient with the larger absolute value at each location in the wavelet domain. This coefficient corresponds to the sample with higher activity level as it represents the most dominant features at each scale in the source images, such as edges, lines and region boundaries. It is defined as:

$$c_{F_n}(j,k)= \begin{cases} c_{A_n}(j,k) & \text{if } \left|a_{A_n}(j,k)\right| > \left|a_{B_n}(j,k)\right| \\ c_{B_n}(j,k) & \text{if } \left|a_{B_n}(j,k)\right| > \left|a_{A_n}(j,k)\right| \\ \dfrac{c_{A_n}(j,k)+c_{B_n}(j,k)}{2} & \text{if } \left|a_{B_n}(j,k)\right| = \left|a_{A_n}(j,k)\right|. \end{cases} \tag{3.12}$$

Using the above fusion rule, the dominant features at each scale are preserved in the new multiresolution representation. However, this rule assumes that only one of the source images provides the relevant information at each scale for fusion. This might not be true, especially when multimodal sensors are used.

Fusion Rule 2 – Weighted average of modes (pixel based)

A second approach based on a weighted combination of the source images is proposed in [10]. The matching measure is used to determine the respective contribution by the different source images. It is given by:

$$c_{F_n}(j,k)= \begin{cases} wc_{A_n}(j,k)+(1-w)c_{B_n}(j,k) \text{ if } \left|a_{A_n}(j,k)\right| > \left|a_{B_n}(j,k)\right| \text{ and } m_{F_n}(j,k) \le T \\ wc_{B_n}(j,k)+(1-w)c_{A_n}(j,k) \text{ if } \left|a_{B_n}(j,k)\right| > \left|a_{A_n}(j,k)\right| \text{ and } m_{F_n}(j,k) \le T \\ \dfrac{c_{A_n}(j,k)+c_{B_n}(j,k)}{2} \text{ if } m_{F_n}(j,k) \ge T, \end{cases} \tag{3.13}$$

where w is the weighted value defining the contribution of the selected coefficient and T is a pre-defined threshold. The larger weight w is assigned to the input image with higher activity level, and can take a range of values from 0.5 to 1.

The fused coefficient corresponds to a weighted average of the input coefficients at each location if the corresponding coefficients in the multimodal images are distinctly different ($m_{F_n}(j,k)$ less than a defined threshold T). If they are similar ($m_{F_n}(j,k)$ greater than a defined threshold T), the average of the two input coefficients will be taken. In the present framework, a value of 0.8 is selected. This can be changed by considering the functional relationship between the weights, activity measure and salience match measures.

Fusion Rule 3 – Weighted average of window-based modes

In the next approach, the scheme takes into account the neighborhood of the selected coefficient. In this fusion scheme, the window-based activity measure $a_{W(A,B_n)}(j,k)$ from Equation (10) replaces the activity measure $a_{A,B_n}(j,k)$ in Equation (13). The fusion rule is given as:

$$c_{F_n}(j,k)= \begin{cases} wc_{A_n}(j,k)+(1-w)c_{B_n}(j,k) \text{ if } \left|a_{W(A_n)}(j,k)\right| > \left|a_{W(B_n)}(j,k)\right| \text{ and } m_{F_n}(j,k) \le T \\ wc_{B_n}(j,k)+(1-w)c_{A_n}(j,k) \text{ if } \left|a_{W(B_n)}(j,k)\right| > \left|a_{W(A_n)}(j,k)\right| \text{ and } m_{F_n}(j,k) \le T \\ \dfrac{c_{A_n}(j,k)+c_{B_n}(j,k)}{2} \text{ if } m_{F_n}(j,k) \ge T. \end{cases} \quad (3.14)$$

In this section, the framework to develop the wavelet-based fusion is presented. First, the source images are decomposed into the corresponding approximate and detail wavelet coefficients. Next, different fusion rules are implemented to determine the relative contribution of the source images. An inverse wavelet transform of the composite wavelet coefficient map produces the fused image.

Other fusion rules and approaches have been proposed using the wavelet-based fusion techniques. Similarly, different wavelet basis functions and a variation to the number of stages of wavelet decomposition can be explored. It is anticipated that some wavelets will be more effective than others and the sharpness of the fused image may improve up to a certain optimum level of decomposition. In this thesis, Daubechies wavelets and up to three levels of decomposition are implemented. It is not possible to consider and implement other configurations within the scope of this thesis; therefore the intent is to lay down the framework of development so interested parties can follow up with the studies.

The next section presents fused results obtained from different image pairs using different fusion rules. It also compares the results achieved when the wavelet transform parameters are varied.

2. Experimental Results – Wavelet Transform Fusion

This section presents the experimental results obtained using wavelet transform fusion. In [12], Nikolov et al. noted that the quantitative measurements of the fused results determined using computational measures are often meaningless or even misleading; therefore the evaluation of the fusion results will

be based on a perceptual comparison of the resultant image with the original images. The assessment will be based on key criteria such as contrast, edge sharpness and scene content.

Test 4-1

Test Objectives:	To demonstrate image fusion using wavelet transform fusion on a pair of out-of-focus images and compare the results achieved with the simple averaging method.
Levels of Decomposition:	2 levels
Wavelet family:	Daubechies, db2
Fusion scheme:	Fusion Rule 1 – Selection of the dominant mode

Figure 19 shows two registered images of the same scene, but with a distribution of defocus. Also shown are their wavelet transforms, the fused wavelet transform and the resulting fused image. The implemented fusion rule - selection of the dominant mode, picks the “detail” coefficients with the largest magnitude at each level. This effectively retains the ‘in focus’ regions within the image. An inverse wavelet transform is then applied to the combined wavelet coefficients to produce the fused image. Figure 19 shows an image retaining the focused regions from each of the two source images.

Figure 20 compares images fused by simple averaging and wavelet transform methods with the original image. In the simple averaging method, the fused image has a “muddy” appearance. A closer inspection of the images shows that the contrast of the features, e.g., roofline, in the fused image is reduced by the averaging process. This results in the blurring of the texture information. Such effects are undesirable in the fusion of night scene images used in applications like night piloting for navigation and target discrimination. Conversely, the multiscale fused approach preserves the texture information and has very good feature contrast. The reconstructed image closely resembles the original image.

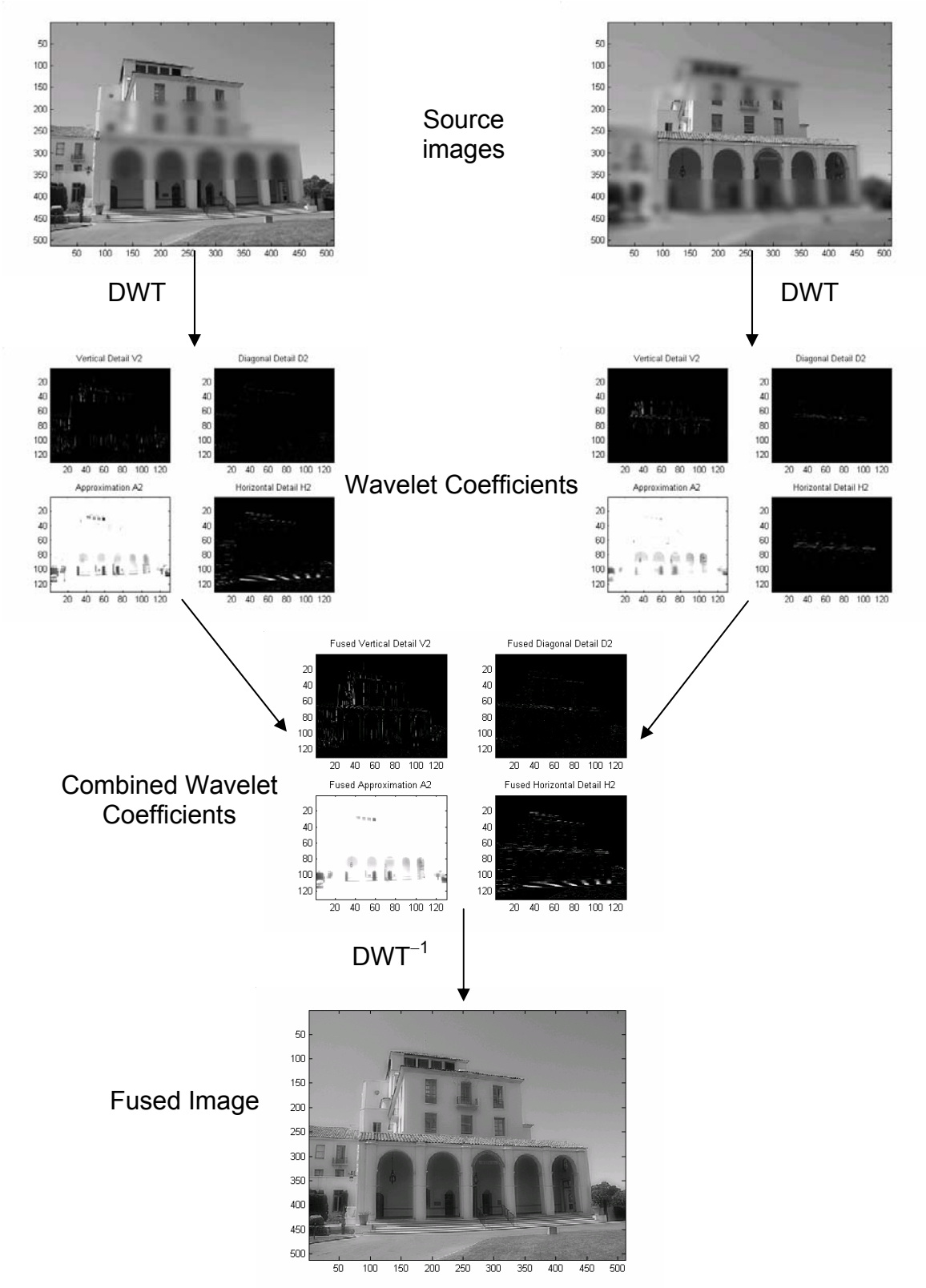


Figure 19. Image fusion process using DWT on two registered multifocus images.

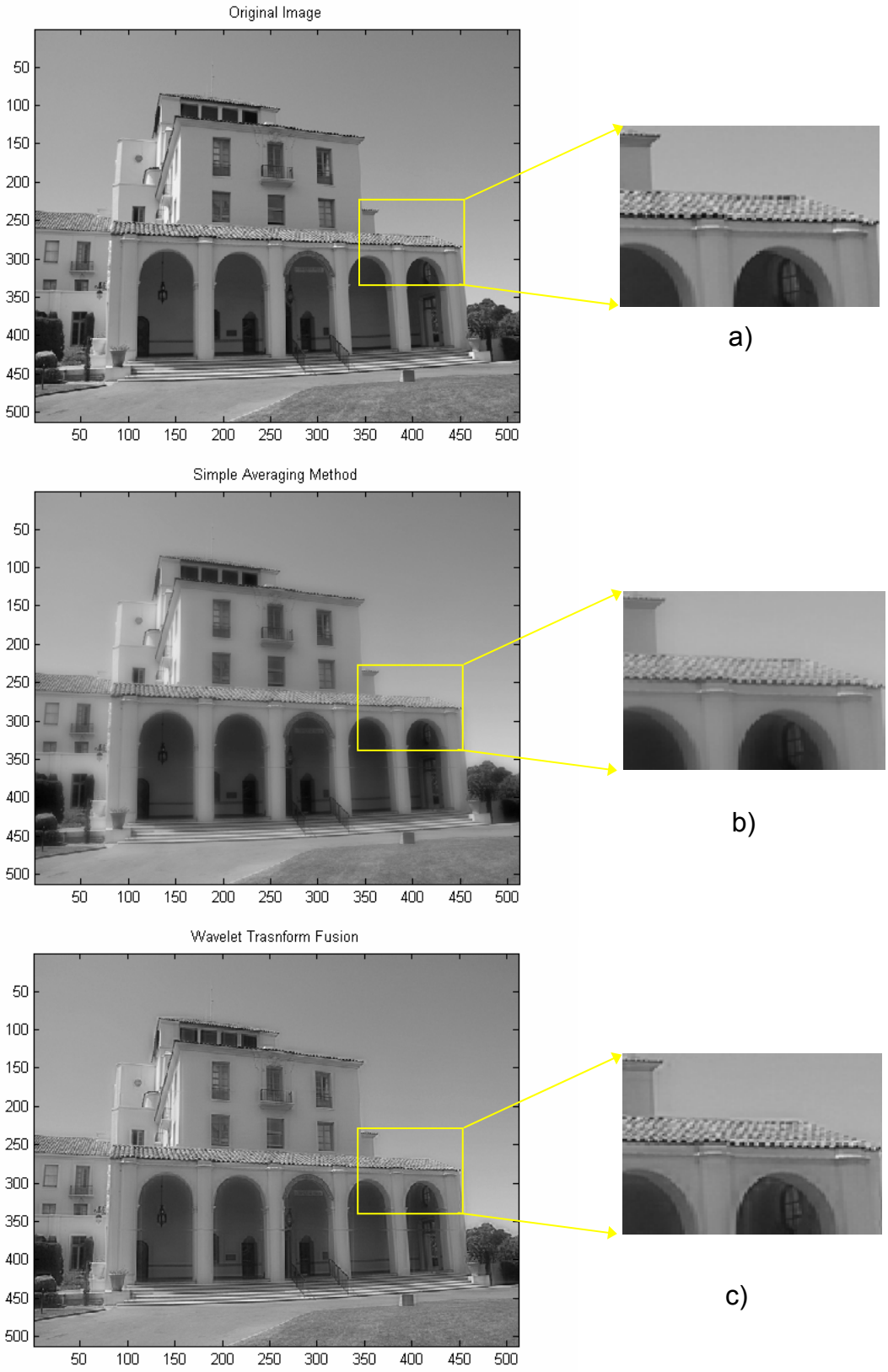


Figure 20. Comparison between simple averaging method and wavelet transform fusion, a) original image, b) fusion using simple averaging and c) wavelet transform fusion using fusion rule 1. The high spectral information in the roofline is retained using wavelet transform fusion.

Test 4-2

Test Objectives:	To implement and evaluate the performance of wavelet transform fusion on a pair of NVD and thermal IR images. The results achieved using 2 and 3 decomposition levels are also compared.
Levels of Decomposition:	2 and 3 levels
Wavelet family:	Daubechies, db2
Fusion scheme:	Fusion rule 1

Figure 21 shows a pair of NVD and thermal IR images of the same scene that were fused using the wavelet transform approach with fusion rule 1. Note that each source image shows certain aspects of the scene that are not visible in the other source. In the fused image, the salient features of the source images are retained. The treeline which divides the image into the top and bottom regions, and the two bright artificial light sources from the NVD image are clearly reflected in the fused image. Similarly, the texture in the foreground, including the track and its adjacent terrain are filled in correctly with inputs from the thermal IR image. The information presented in the fused image is much richer than that contained in either source image and would be essential for situation awareness and navigation.

The fused images obtained using 2 and 3 decomposition levels are displayed in Figure 21. The inset in (b) (3 levels) shows greater contrast and “graininess” than the corresponding inset in (a), which presents a more pleasing picture. With a higher level of decomposition, features found only in the coarser scale are also extracted using the dominant mode selection rule. Therefore, the result is a fused image that has a slightly better spectral quality. However, it is not recommended to go beyond 3 levels of decomposition as the loss of details of the approximate sub-image increases with the number of decomposition layers and reconstructing the lost details would be difficult [20].

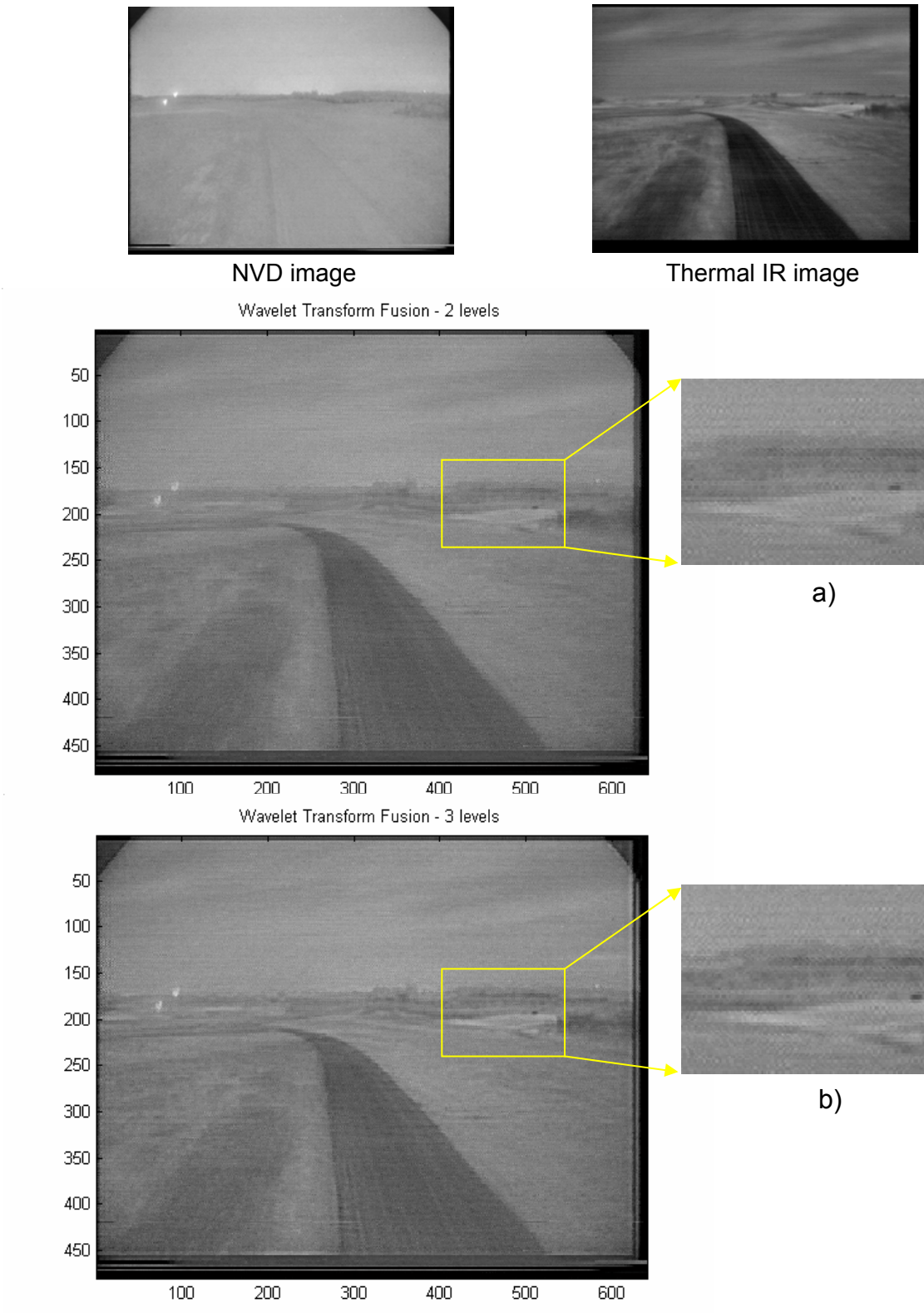


Figure 21. Fusion of NVD and thermal IR images with a) 2 levels and b) 3 levels of decomposition, using fusion rule 1(source images from Naval Research Laboratory).

Test 4-3

Test Objectives:	To implement and compare the performance of wavelet transform fusion on a pair of NVD and thermal IR images, using fusion rule 1 - selection of the dominant mode, fusion rule 2 - weighted average of modes (pixel based) and fusion rule 3 - weighted average of window-based modes.
Levels of Decomposition:	3 levels
Wavelet family:	Daubechies, db2
Fusion scheme:	Fusion rule 1, 2 and 3

Figure 22 shows the results achieved using the three different fusion rule schemes. In fusion rule 3, a small neighborhood consisting of 3 by 3 arrays of samples centered on sample was used to compute its windowed-based activity measure. All three cases generate a perceptually similar fused image. The feature contrast is well maintained and all the significant features from both sources, e.g., the two artificial light sources, night sky and the track, are retained in the composite image.

It is noted during the test that the relative contribution of the source images to the fused image can be changed by varying the weighting factor and matching threshold. This will alter the spectral contrast of the resulting fused image.

In summary, experimental results show that the wavelet-based approach outperforms the simple averaging method and offers significant scene content improvement over single sensor detection. Different fusion rule schemes have been implemented and they perform well in the fusion of the NVD and thermal IR images. The choice of the fusion rule scheme as well as the selection of the weighting factor and matching threshold will be application specific and is likely to depend on the type of image sensors, scene composition, target types etc. The functional relation between the fusion rule scheme, weighting factor, activity measure and salience match measures can take many forms and further tests and evaluations are needed to determine the optimal configuration.

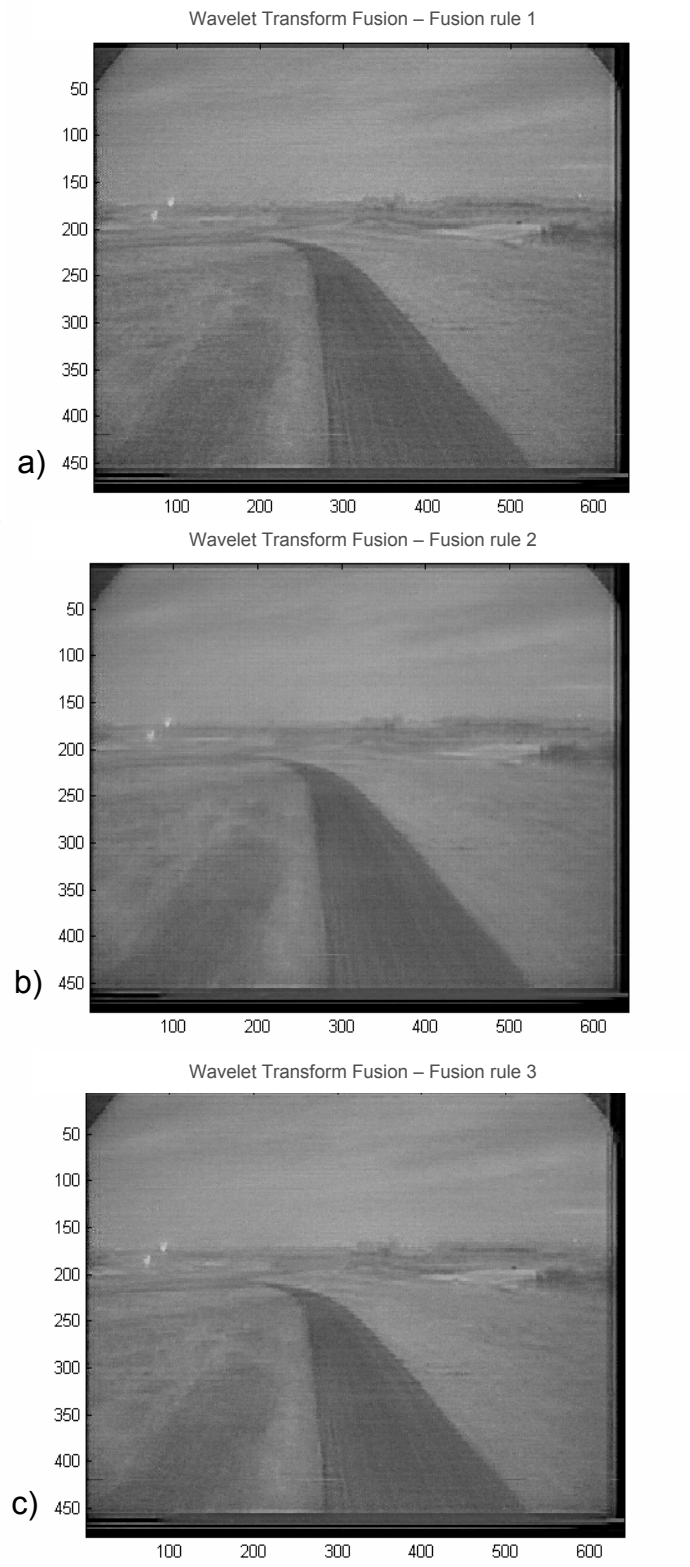


Figure 22. Fusion results achieved with using 3 levels of decomposition a) fusion rule 1- selection of the dominant mode, b) fusion rule 2 - weighted average of modes (pixel based) and c) fusion rule 3 - weighted average of window-based modes.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. REGION-BASED FUSION

A. OVERVIEW

Fusion methods based on relatively simple image processing techniques, e.g., the pixel-level averaging method, generally do not take into account the subject-relevant information or the features that exist in the source images. If the features information is not incorporated in the fusion process, it could lead to undesirable effects such as artifacts or inconsistencies and the loss of vital information in the fused image. In the previous chapter, a wavelet based pixel-level fusion method which combines aspects of a feature selection rule was implemented. The fusion process is guided by the salient directional features identified at the multiscale detail images. This is done by comparing the intensity of the corresponding pixels or an arbitrary area around the sampled pixel defined by a fixed size window in the corresponding detail images, and selecting one deemed more important for the fused pyramid. Experimental results show that the algorithm works well in fusing image pairs captured by NVD and thermal IR sensors.

To increase the degree of subject relevance in the fusion process, region-based fusion schemes have been proposed [18,19,20,21 etc]. They are based on segmenting the multimodal source images into regions of interest and subsequently using this segmentation to guide the fusion process. Region-based image fusion algorithms are known to be more robust and less sensitive to noise and misregistration. A number of different region-based schemes have been proposed. In [21], a Canny edge detection method was applied to the approximate sub-image obtained from the wavelet transform. This edge information is then used to obtain the segmentation of the low frequency band. In [18], the author proposed a region-based MR fusion scheme using a segmentation algorithm based on a generalized pyramid linking method.

In this thesis, a segmentation algorithm based on the watershed transform is investigated. It is combined with the results derived using the wavelet transform in Chapter III to implement a region-based fusion scheme. By incorporating the region information, the proposed approach seeks to optimally extract the information from different sources and maximize the "scene content" in the fused image. The following topics will be covered in this chapter: 1) implementation of the watershed transform for image segmentation; 2) an investigation of multiscale image segmentation, and 3) the theory and experimental results of region-based image fusion.

B. REGION SEGMENTATION

The objective of segmentation is to partition an image into a number of disjoint regions in each of which the features should have reasonably good homogeneity, strong statistical correlation or visual similarities. Image segmentation algorithms may be generally classified into discontinuity-based methods and similarity based methods [22]. The interface between two homogenous regions is usually defined by a discontinuity in gray-level, color or texture. Discontinuity based methods therefore partition an image based on the detection of such discontinuity (gradient). Segmentation based on the similarity method typically works by detecting homogeneity between pixels and regions, and the image is segmented according to certain pre-defined criteria or levels. Each approach has its own pros and cons in terms of applicability, performance and computational cost etc. A good guideline defining segmentation is given in [23]. It stated the following requirements: "1) Regions of an image segmentation should be uniform and homogeneous with respect to some characteristic such as gray tone or texture; 2) Region interiors should be simple and without many small holes; 3) Adjacent regions of a segmentation should have significantly different values with respect to the characteristics on which they are uniform, and 4) Boundaries of each segment should be simple, not ragged and must be spatially accurate."

The classic approach to segment an image is to apply a gradient and then threshold the resulting gradient image. However, it is difficult to select an appropriate value for thresholding. If the threshold value is too low, false edges and noise are picked up and lead to inaccurate segmentation. Conversely, edges may not be detected if the threshold is too high. As a result, broken gradients would form and result in poor segmentation.

An alternative method based on morphological principles, watershed transformation, has evolved and become a well established approach for the segmentation of images. Mathematical morphology is a nonlinear image processing and analysis tool that describes the basic characteristics of an image, namely the geometry and structure relation between the pixel sets in the image using a set of integrated concepts and algorithms. It uses a structuring element with a certain shape to measure and detect objects with a corresponding shape in the image. By marking the location where the structure fits, the structural information in the image can be derived [24].

Instead of using the image directly, the watershed transform algorithm is applied to the morphological gradient of the image to be segmented. Implementations of the watershed approach on the test images yielded promising results and, therefore, it will be used to identify the key regions in the multimodal source images during pre-processing prior to the fusion of images. The following section presents the approach adopted and results achieved using the watershed transformation.

1. Watershed Transform

A grayscale image can be considered as analogous to a topographical relief map with the brightness value of each pixel corresponding to a physical elevation at that point. If this topography is flooded from below, water will slowly rise from each regional minimum at a uniform rate across the entire image. A dam is created when water from two different regions meets. The procedure results in the partitioning of the image in which the different regions arising from

the various regional minima are called the catchment basin [25]. Figure 23 illustrates the principle of the watershed transform.

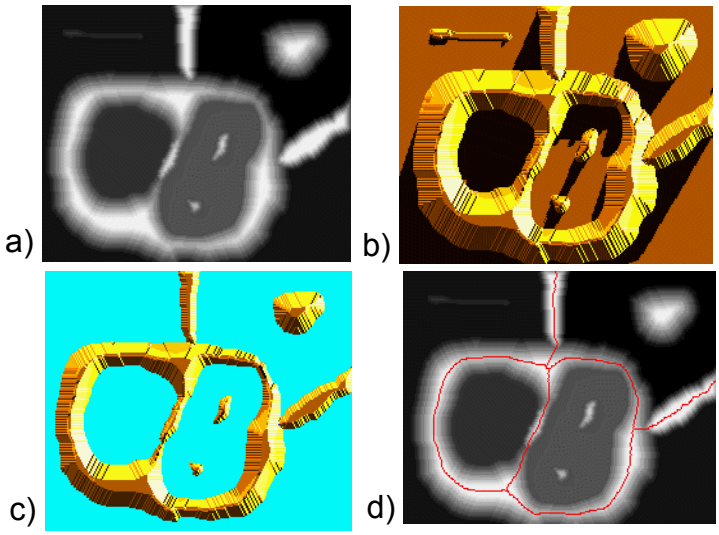


Figure 23. Principle of watershed transform: a) grayscale image; b) topographical surface; c) flooding in the basins; d) watershed (From Ref. [25]).

The watershed transform is applied to a gradient image so that the watersheds correspond to the crest line of the gradient. Therefore, the catchment basin maps to the regions in the image. The gradient is created by standard morphological operations, namely “Dilation” and “Erosion”. Following reference [24], the morphological definitions are given as follows. The erosion of the binary image set A by a small set B, representing the structuring element is defined as:

$$A \ominus B = \{ x : B_x \subset A \} , \tag{4.1}$$

where $\subset$ denotes the subset relation, A the input image, B the structuring element and B_x is the translation of B along vector x. $A \ominus B$ consists of all points of x for which the translation of B by x fits inside of A and represents a filtering on the inside. Dilation is the dual operation to erosion and is defined via erosion by set complementation. It is defined by:

$$A \oplus B = (A^c \ominus \breve{B})^c , \tag{4.2}$$

where

$$\breve{B} = \{-b : b \in B\} \tag{4.3}$$

is the reflection of B or a 180-deg rotation of B about the origin and A^c denotes the set-theoretic complement of A. Dilation represents a filtering on the outside A by B.

The morphological gradient is given by the differences between the dilation and erosion and is given by:

$$(A \oplus B) - (A \ominus B), \tag{4.4}$$

Figure 24 illustrates boundaries created using a four-connected structuring element. Geometrically, in erosion, the structuring element B is moved within the image A. The origin of the structure is marked in dark blue and represents the eroded image. In dilation, the origin of the structure is moved along the boundary of the image A. Pixels overlapped by the 4-connected structuring element are combined with the image A to form the dilated image. The morphological gradient is given the difference between dilated and eroded image.

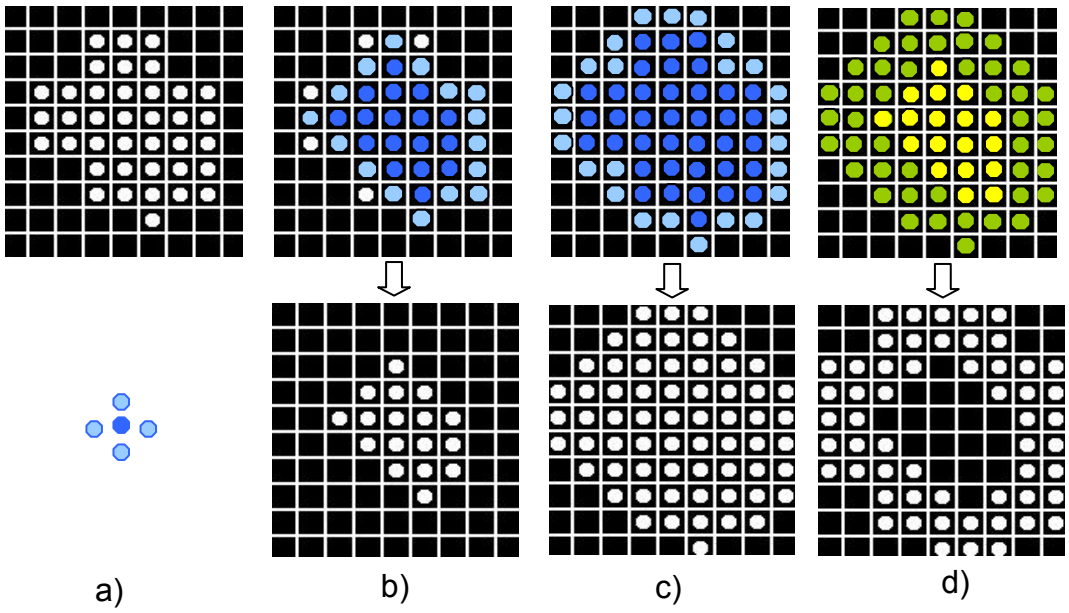


Figure 24. Boundary creation: a) input image, A and a four-connected structuring element, B; b) erosion of A by B, c) dilation of A by B, d) morphological gradient (From Ref. [24]).

Direct application of the watershed transform to a gradient usually produces excessive over-segmentation (Figure 25). This is undesirable as the segmented regions do not offer a good local characterization of the region. Therefore, a marker-based watershed segmentation is implemented.

The marker is a connected component belonging to an image and it guides the flooding simulation process, thereby leading to a marked improvement in the segmentation results. The number of regions segmented is reduced as the marker decreases the number of minima on the surface. A marker-based watershed segmentation scheme was implemented[†]. Figure 26 presents the results achieved. It demonstrates the following advantages in image segmentation: a) closed and connected regions are formed, unlike traditional edge based techniques that tend to form disconnected boundaries, b) the boundaries of the resulting regions correspond well to the contours in the images, and c) the union of all the regions forms the entire image region. The advantages highlighted are critical to the successful implementation of the fusion approach proposed in the next section.

[†] The morphological functions are implemented using SDC's Morphology Toolbox for MATLAB (From Ref. [26])

Figure 25. Simple watershed transform – Oversegmentation, showing tile-like structure.

a)

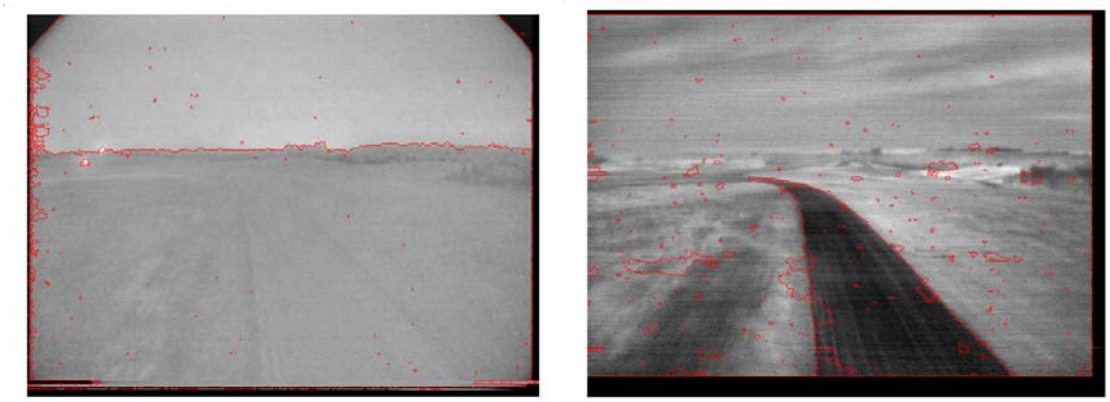
b)

c)

Figure 26. Marker-based watershed segmentation: a) morphological gradient, b) watershed lines overlying the original image and c) identified regions.

2. Multiscale Segmentation of Images

As quoted in [22], “Segmentation of nontrivial images is one of the most difficult tasks in image processing.” Clearly, the raw images of scenes captured by NVD and thermal IR sensors are nontrivial images, as they generally do not have well defined regions that are characterized with good homogeneity or clear boundaries. They tend to have low contrast edges and are noisy. Any noise-induced gray level fluctuations can result in spurious gradient and further complicate the segmentation process. Figure 27 shows the outline of the regions obtained using marker-based watershed segmentation. A number of smaller undesired watersheds are generated and this results in oversegmentation despite using a marker-based approach. As segmentation accuracy determines the eventual success or failure of the next stage of the fusion process, it is necessary that further pre-processing be done to produce a segmentation that better identifies the regions in the image.



a) b)

Figure 27. Segmentation using marker-based watershed segmentation on: a) NVD image and b) thermal IR image.

The threshold method used in marker-based watershed segmentation is not sufficient to eliminate undesired gradients. Methods using conventional filtering methods have been explored and implemented to reduce the small details in the image, e.g. gradient caused by noise or other minor structures. However, the results are generally less than satisfactory in complex images when low contrast edges are involved or in high noise level.

In [27], Jung et al. proposed a wavelet-based approach to denoise and to enhance the edges of the image. The watershed transform is then applied to the gradients of the enhanced image to segment the image. A final post-processing is done to remove the regions with small areas and to merge regions with low contrast boundaries. Preliminary results show that oversegmentation is reduced and broken contours are significantly removed.

The test images used by Jung et al. consisted of regions of cluttered objects that are relatively homogenous. In this thesis, a concept similar to [27] is proposed. The new approach combines the multiscale wavelet transform introduced in Chapter III with the morphological watershed transform to segment the image with the objective of generating a well segmented image that can be used to guide the fusion of the multimodal images. It will be applied to images captured by NVD and thermal IR sensors and the test will be challenging as they tend to have regions of non-uniform homogeneity, low contrast and poorly defined boundaries (Refer to Figure 5 and Figure 6).

In accordance with Equation (3.7) and Figure 11 in Chapter III, a source image can be decomposed into an approximation a_{LL_1} and three detail images, c_{LH_1}, c_{HL_1} and c_{HH_1} at every level of decomposition. The approximate sub-image represents the averaged, lower resolution, version of the base low frequency image from the previous level while the details images captures the local differences or texture along the horizontal, vertical and diagonal fluctuations in that image.

To improve the performance of the segmentation, the watershed transform is applied to the approximate sub-image at every level of decomposition. Since the n^{th} level approximate sub-image contains less detail than the $(n-j)^{th}$ level approximate sub-image, the reduction in detail would improve the quality of the segmentation based on the watershed transform. The idea is similar to the application of the wavelet transform for image denoising where the wavelet coefficients in the detail images correspond to the high frequency components at that scale. Therefore, by applying an appropriate threshold to these coefficients,

that is, setting coefficients to zero whose magnitude is less than the threshold value, the inverse transform of the thresholded transform reduces the noise level of the original source image.

Using the algorithms generated in Chapter III and the earlier sections in Chapter IV, marker-based watershed segmentation is applied to the morphological gradient of the approximate image at every level to extract the various regions at each scale.

The above segmentation procedure is applied to the NVD and thermal IR image pair (Figure 5 and Figure 6).The morphological gradient operator is first applied to the coarse approximates of the NVD and thermal IR images; the gradient of the pixel values is then plotted over the source images. In this image, uniform regions with large gradient (greater than threshold) are partitioned using the marker-based watershed segmentation technique and they show as topographical relief features. Results are shown in Figure 28 to Figure 31.

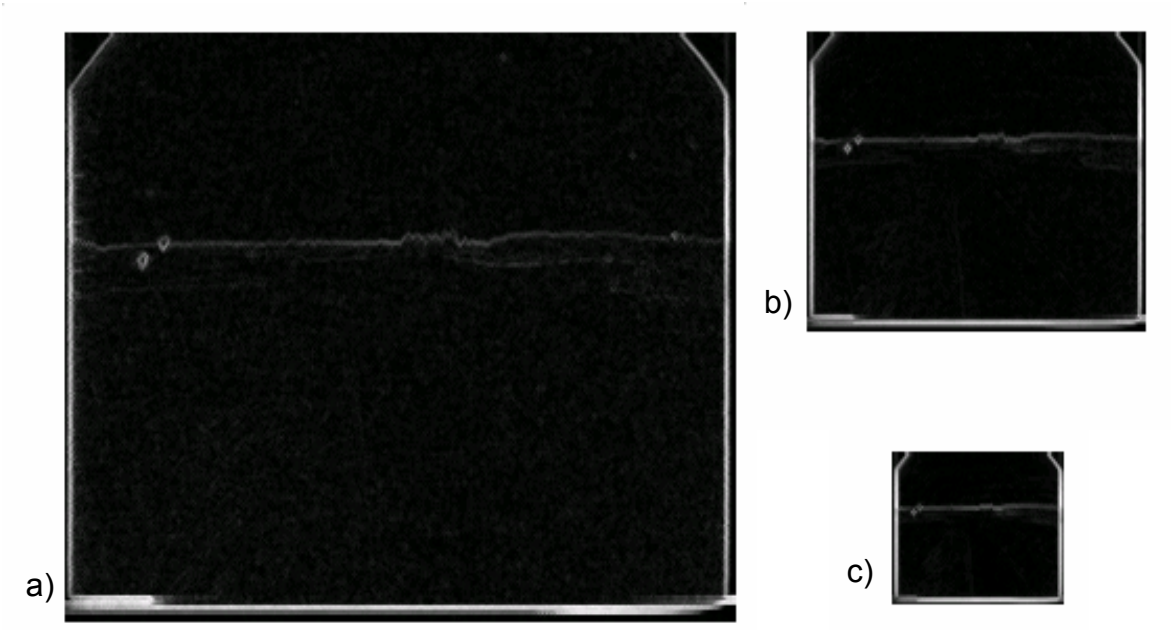


Figure 28. Morphological gradient of the approximate NVD image at different 3 levels of decomposition: a) level 1, b) level 2 and c) level 3.

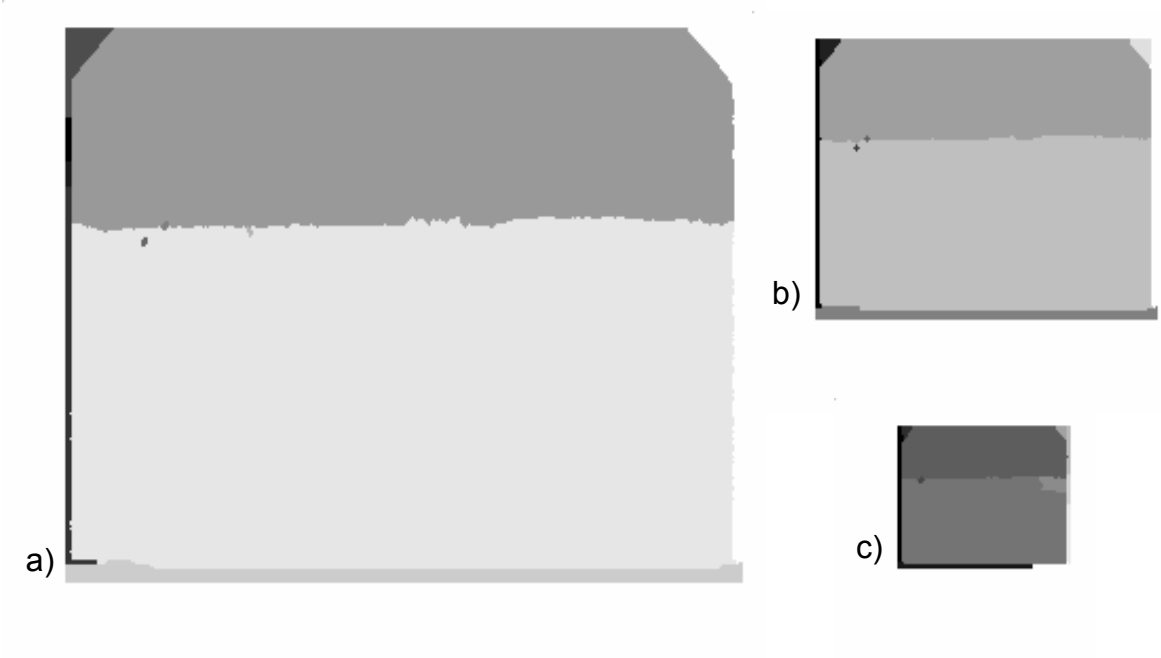


Figure 29. Region segmentation of the approximate NVD image at 3 levels of decomposition: a) level 1, b) level 2 and c) level 3.

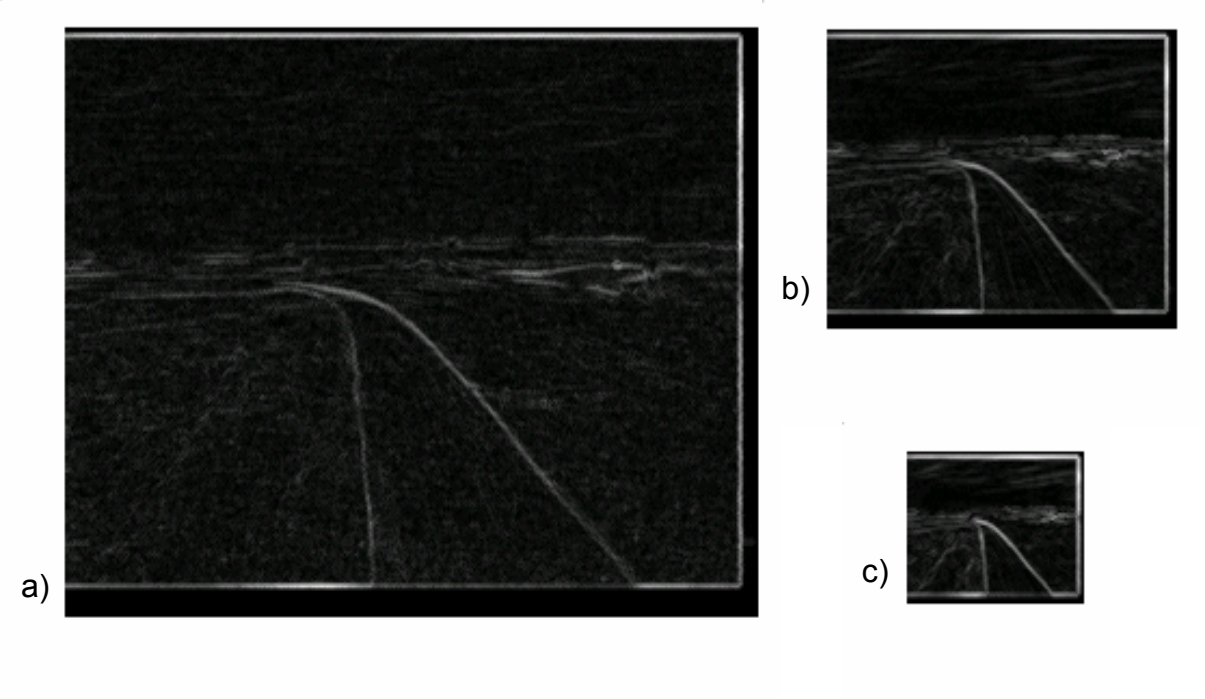


Figure 30. Morphological gradient of the approximate thermal IR image at 3 levels of decomposition: a) level 1, b) level 2 and c) level 3.

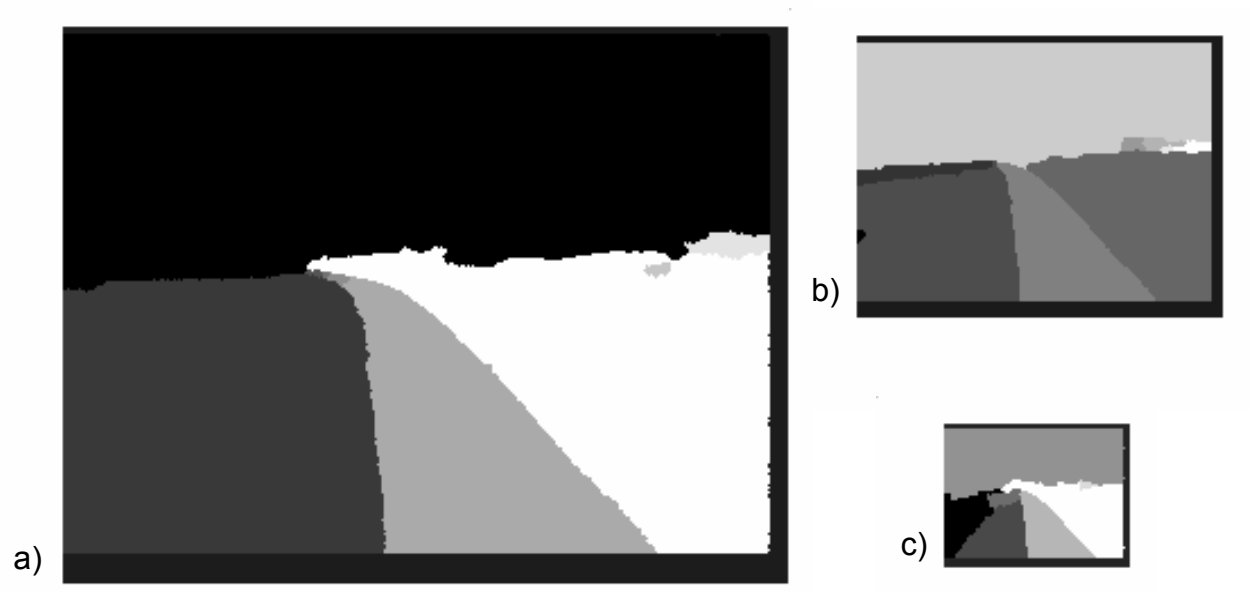


Figure 31. Region segmentation of the approximate thermal IR image at 3 levels of decomposition: a) level 1, b) level 2 and c) level 3.

Comparing the above results to Figure 27, it is clear that the regions are more accurately segmented without leading to oversegmentation. An interesting observation is that the feature edges are preserved in the lower scaled sub-images. This is to be expected as the responses due to noise tend to be more localized and therefore are not likely to be present across the different scales.

The watershed transform of the approximate images is guided by setting the number of regions in the joint region map. Results show the process generally produces a good segmentation of the test images by limiting to under forty regions. The computation time for the subsequent stages in the fusion process increases with the number of regions segmented; therefore limiting the number of segmented regions also serves to cap the computation time to an acceptable level.

Further post-processing can be done to remove over-segmented regions by merging small watershed regions resulting from weak borders that may still exist in the approximate image [27]. The results achieved here are generally satisfactory; therefore the post processing algorithm is not implemented. However, this step will need to be considered when multi-modal images are fused using region-based techniques.

C. REGION-BASED IMAGE FUSION

The basic idea behind the proposed region-based image fusion is to construct a multiscale segmentation based on the approximate sub-images and to use this segmentation to guide the fusion process. The general framework of the region-based image fusion scheme proposed in this thesis is an extension of that proposed for wavelet transform fusion in Chapter III (Figure 16 and Figure 17). Figure 32 shows the schematic representation of the process of region-based fusion using the fusion rules to be discussed in the section following.

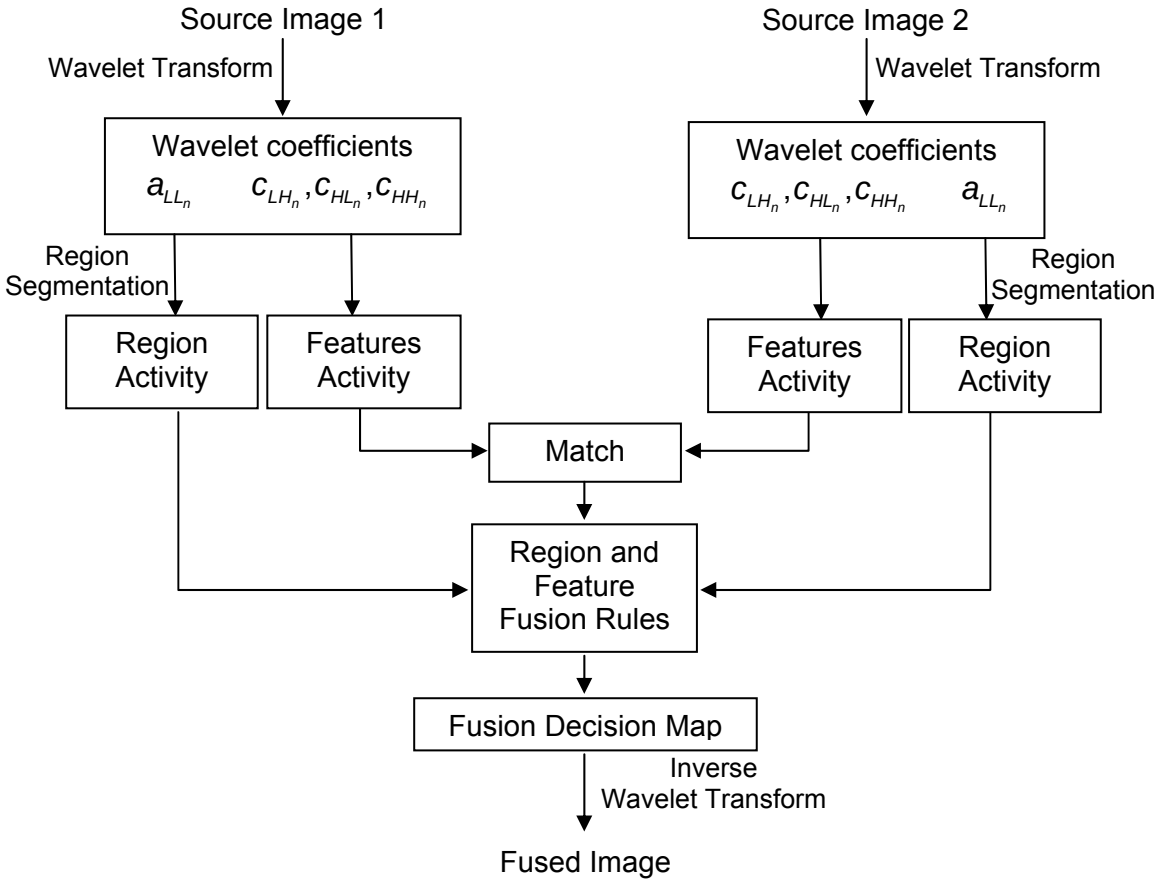


Figure 32. Framework for the formation of the fusion decision map for region-based fusion. It illustrates the process of constructing a "decision map" for region-based wavelet transform fusion of images.

In addition to using a feature selection fusion rule to construct the detail sub-images decision map, a region activity table is generated based on the regions identified on the coarse approximation image using the watershed transform. The region and feature fusion rule is then applied to the corresponding activity table to generate the fusion decision map that will decide how the multiscale representations will be used to construct the fused wavelet coefficient map.

1. Fusion Rules

In the previous section, a multiresolution segmentation performed on the NVD and thermal IR source images produces two region representations R_{A_n} and

R_{B_n} , as shown in Figure 29 and Figure 31. To identify all the regions in the source images, the two region representations are overlaid onto each other to create a joint region map R_{F_n} at each level of decomposition [28]. The concept is illustrated below in Figure 33.

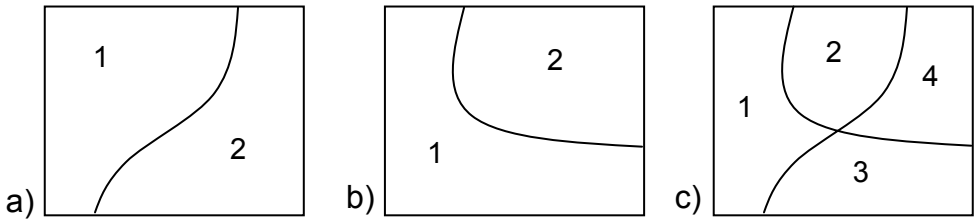


Figure 33. Region segmentation: a) region representation of image A; b) region representation of image B and c) joint region map, indicating the 4 identified regions (After Ref. [28]).

Applying this concept to NVD and thermal IR source images, the joint region maps obtained at different levels are shown in Figure 34. The disjoint regions corresponding to unique features of the two image sets are combined together and will be used to guide the computation of the activity level of each region in the decomposed approximate sub-images.

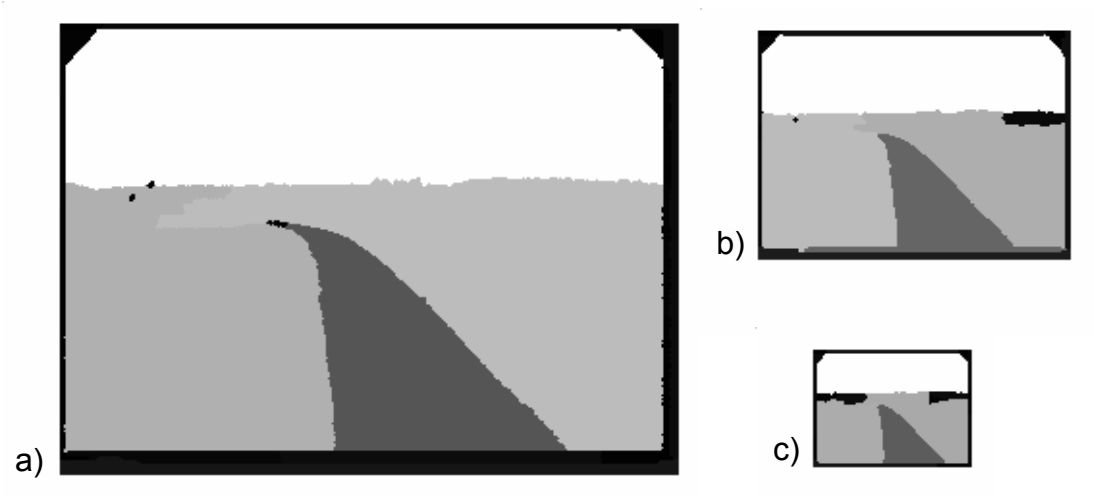


Figure 34. Joint region maps for NVD and thermal IR images at different levels of decomposition, a) level 1, b) level 2 and c) level 3.

To compute the region activity, the following steps are implemented.

Step1: The regions identified in the multiscale joint region maps are assigned a label,

$$R = \{R^k{}_n\}, \tag{4.5}$$

where $R^k{}_n$ represents the k^{th} segmentation at level n. This label will be used to mark and identify the pixels lying within the boundary of a region.

Step2: Determine the size of the regions. This is given by the total number of pixels within the boundary of the region. The joint region map for the NVD and thermal IR images is illustrated in Figure 35. It shows the size of the two artificial light sources relative to the foreground terrain and night sky background.

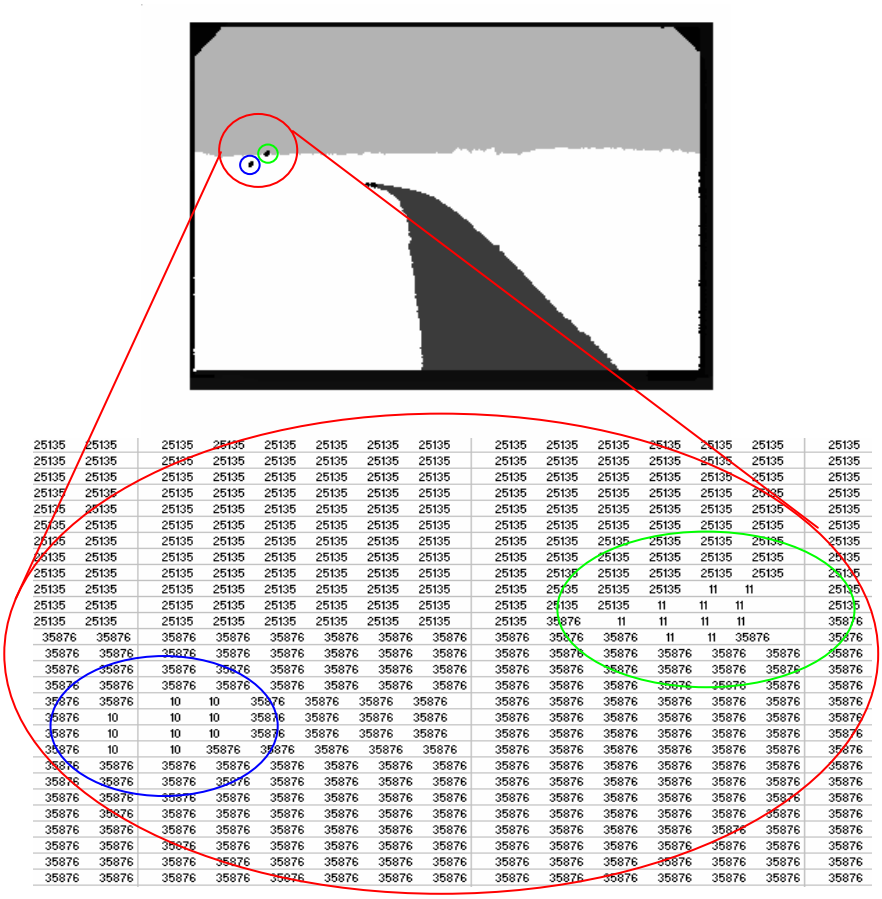


Figure 35. Illustration of the computed region size in the joint region map. The large elliptical region (red) contains 35867 pixels while the two artificial light sources, shown as the small elliptical insets (blue and green) contain 10 and 11 pixels respectively.

Step3: Overlay the boundaries of the joint region map onto the source images. This allows a visual inspection of the region activity level for the respective source images in the joint region maps, as shown in Figure 36 and Figure 37.

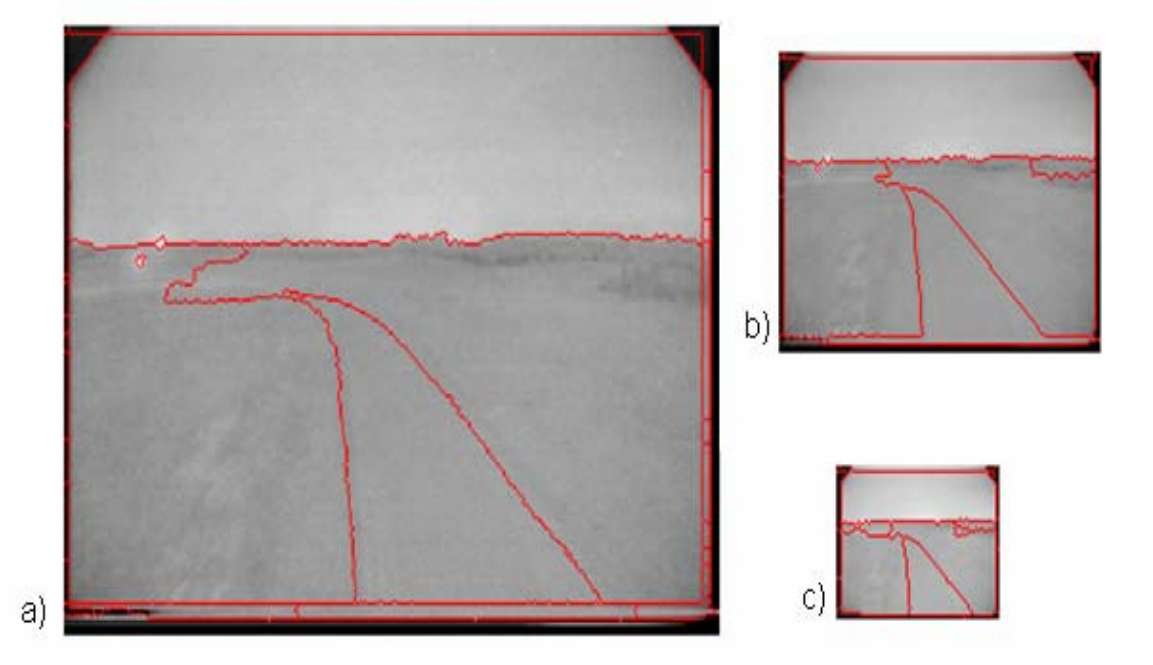


Figure 36. Boundaries of the joint region map are plotted over the NVD source image, highlighting the outstanding features present in this image, e.g., artificial light sources, background night sky and foreground terrain.

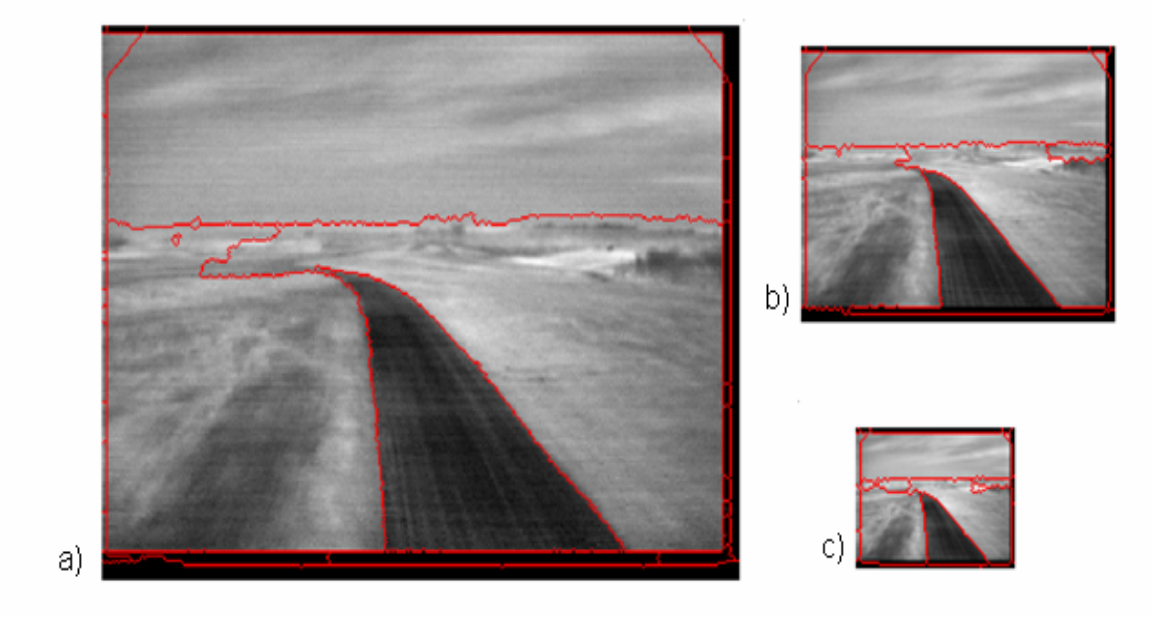


Figure 37. Boundaries of the joint region map are plotted over the thermal IR source image, highlighting the outstanding features present in this image, e.g., track and foreground terrain texture.

Step4: Compute the activity measure of each region for both the source images. The activity level of region *x* in the image *A*, is given by:

$$A_{A_n}(x)=\frac{1}{S_i}\sum_{S_i} a_{A_n}(j,k), \tag{4.6}$$

where $a_{A_n}(j,k)$ is given by Equation (3.9) and represents the n^{th} level activity measure of the wavelet coefficients at location (*j,k*), S_j is the size of the region determined in Step 2. This step is repeated for image *B*.

The above information is then integrated to generate a fusion decision map which governs the combination of the coefficients of the transformed sources. In the decision process, the following weighted average fusion rule is implemented for the approximate sub-image at each *n* level and for each region $R^k_n \in R$.

$$c_{F_n}(j,k)=\begin{cases} wc_{A_n}(j,k)+(1-w)c_{B_n}(j,k) \text{ if } \left|A_{A_n}(x)\right|>T \\ wc_{B_n}(j,k)+(1-w)c_{A_n}(j,k) \text{ if } \left|A_{B_n}(x)\right|>T \\ \dfrac{c_{A_n}(j,k)+c_{B_n}(j,k)}{2} \text{ otherwise,} \end{cases} \tag{4.7}$$

where *T* is a threshold defined to identify regions of high activity, *w* is a weighting factor, $c_{F_n}(j,k)$ represents the composite coefficients, and $c_{A_n}(j,k)$ and $c_{B_n}(j,k)$ are the source coefficients of images *A* and *B* respectively. According to the above fusion rule, the composite approximation image is formed by a selective combination of the source image coefficients which are given a weighting corresponding to each region's activity measure. If the regions exhibit similar activity level, the composite coefficients will take the average of the two source coefficients.

In the last two sections, the concept of image segmentation using the watershed transform is discussed. The framework to implement the region-based fusion is then presented. First, the source images are decomposed using the wavelet transform. Next, a marker-based watershed transform is applied to the coarse approximate sub-image to partition it into “regions of interest”. Lastly, the region activity measure is derived and used to guide the fusion of the approximate wavelet coefficients. An inverse wavelet transform on this composite approximate and detail wavelet coefficient map produces the fused image. In the next section, the proposed algorithm is tested on different sets of NVD and thermal IR image pairs.

2. Experimental Results – Region Based Fusion

This section presents the experimental results obtained using the proposed region-based fusion algorithm. The fused images will be evaluated through visual inspection using the key assessment criteria: contrast, edge sharpness and scene content.

Test 5-1

Test Objectives:	To implement and evaluate the performance of the proposed region-based fusion algorithm on a pair of NVD and thermal IR images. The results obtained using different fusion schemes are compared.
Levels of Decomposition:	2 levels
Wavelet family:	Daubechies, db2
Fusion scheme:	Region-based fusion rule

Figure 38 shows the fusion of the NVD and thermal IR source images using the proposed region-based fusion algorithm. The joint region map obtained from the watershed transform is used to derive the decision maps for the approximate sub-images. According to the fusion rule, Equation (4.7), a region having an activity level above the defined threshold, is given a higher weighting in the fusion process. Therefore, the regions corresponding to the road and the two artificial light sources are selected from the thermal IR image and the NVD

image respectively. Most of the background is selected by averaging the coefficients from the two source images. The fused approximate sub-image is shown in Figure 38. Combining with the feature fusion rule, Equation (3.12), for the selection of the detail coefficients, the composite wavelet coefficients are obtained. An inverse wavelet transform applied to these combined wavelet coefficients produces the fused image as shown in Figure 38.

In addition to retaining the key features and texture information from the source images, the fusion process places a greater emphasis on the 'regions of interest'. Compared to pixel level fusion, the fused results obtained using the region-based approach better reflects the scene content of the source images. It demonstrates the potential of region-based fusion using the proposed algorithm.

Figure 39 shows the comparison between the different weighting schemes. A larger weighting factor increases the emphasis on the high activity regions, e.g. track and artificial light sources. For example, the region representing the track in the foreground has a much higher region activity measure in the thermal IR image than the NVD image. Therefore, the larger weighting factor increases the relative contribution of the thermal IR image to the fused image, which leads to better retention of the salient features.

At $w = 0.5$, the fused approximate wavelet coefficient map is obtained by taking the average of the source images' approximate coefficients and the fused results obtained would be the same as that derived using the pixel level wavelet transform fusion.

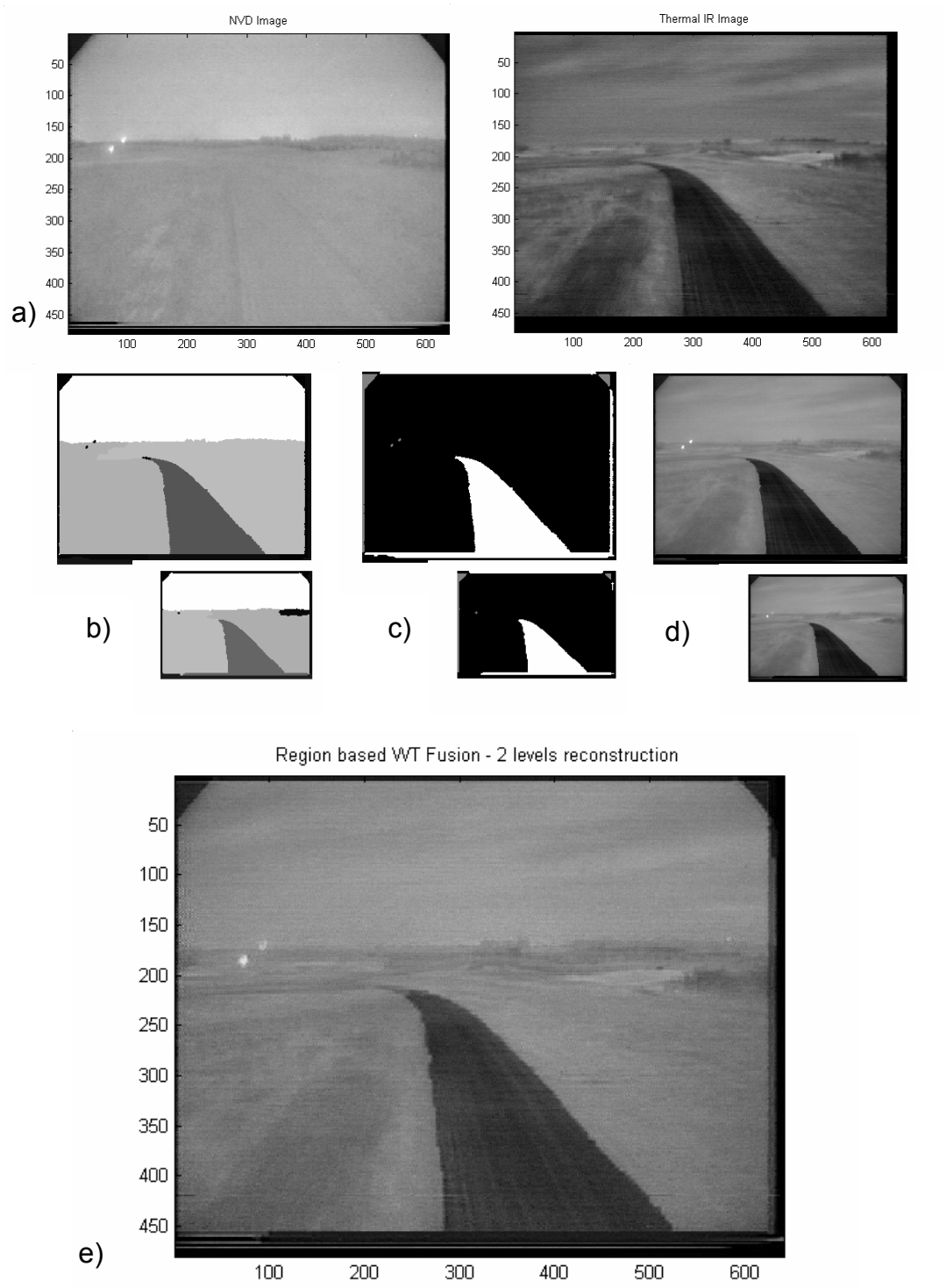


Figure 38. Test 5-1: a) NVD and thermal IR source images; b) Joint region maps achieved using watershed transform; c) level 1 and 2 decision maps; d) level 1 and 2 fused approximate sub-images and e) reconstructed fused image (source images from Naval Research Laboratory).

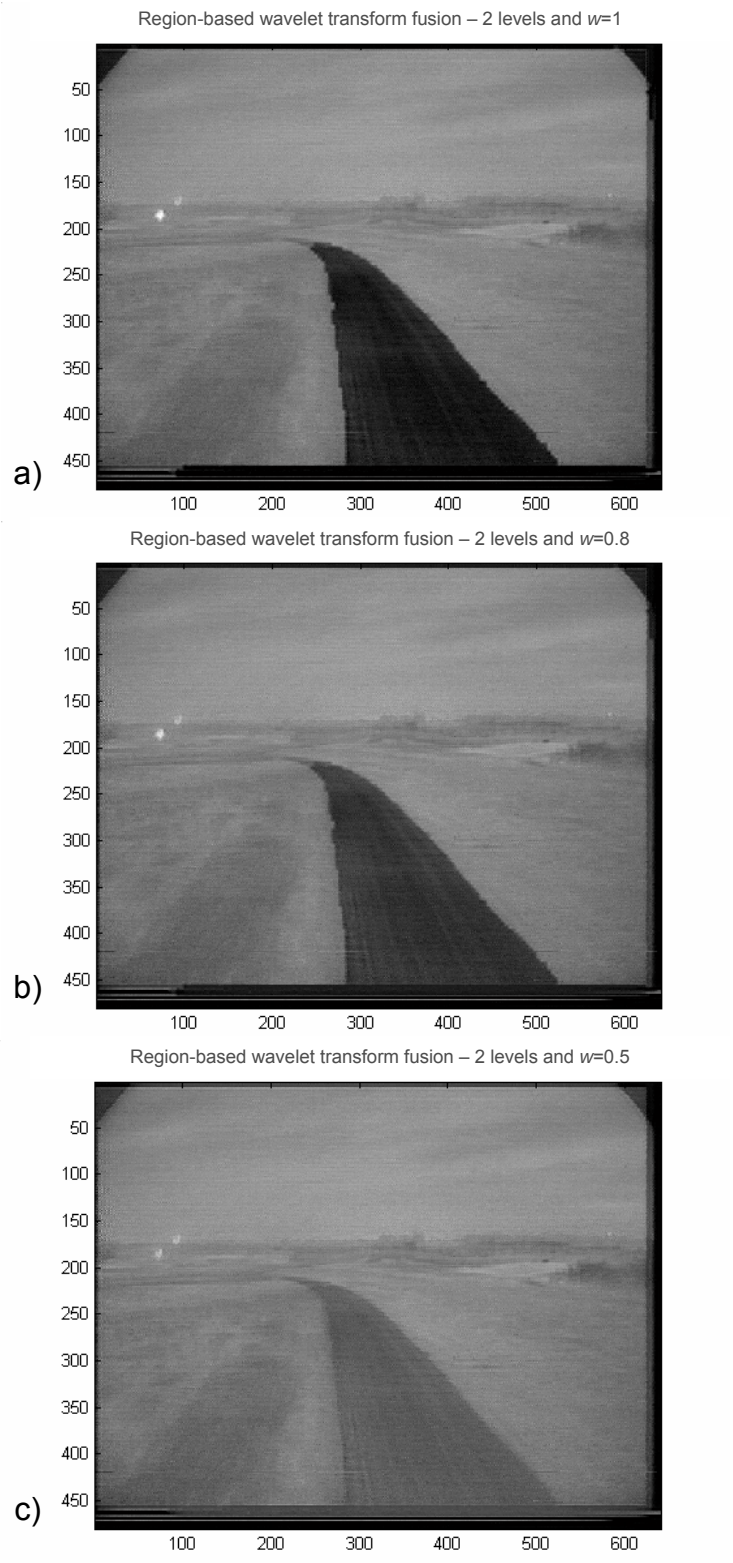


Figure 39. Comparison between different weighting schemes: a) region-based fusion with weighting factor w = 1; b) region-based fusion with weighting factor w = 0.8 and c) region-based fusion with weighting factor w = 0.5, Wavelet transform fusion (pixel level fusion).

Test 5-2

Test Objectives:	To implement and evaluate the performance of the region-based fusion algorithm on a different set of NVD and thermal IR images.
Levels of Decomposition:	2 levels
Wavelet family:	Daubechies, db2
Fusion scheme:	Region-based fusion rule

Figure 40 shows the experimental results of the region-based fusion of a different set of NVD and thermal IR images. The low luminance, coupled to the low reflectivity from the foliage generates a low contrast NVD image that captures limited details of the foreground terrain. The moon and the night sky in the background are more luminous and therefore can be differentiated against the foreground and treeline. The NVD image shows little ‘texture information’. It is complemented by the thermal IR image, which captures the surface details due to the greater contrast in the emissivity of the foreground terrain.

Applying the watershed transform to the approximate sub-images, the source images are partitioned into distinct identifiable regions as shown in Figure 40(b). Except for the region representing the moon in the background, most of the segmented regions do not have a very high activity measure. Thus, the algorithm generates a decision map that emphasizes only the coefficients representing the moon and averages the rest of the coefficients, as shown in Figure 40(c) and Figure 40(d).The final result is presented in Figure 40(e). It shows that the salient features in the respective source images can be emphasized by selecting an appropriate parameter in the region fusion rule scheme.

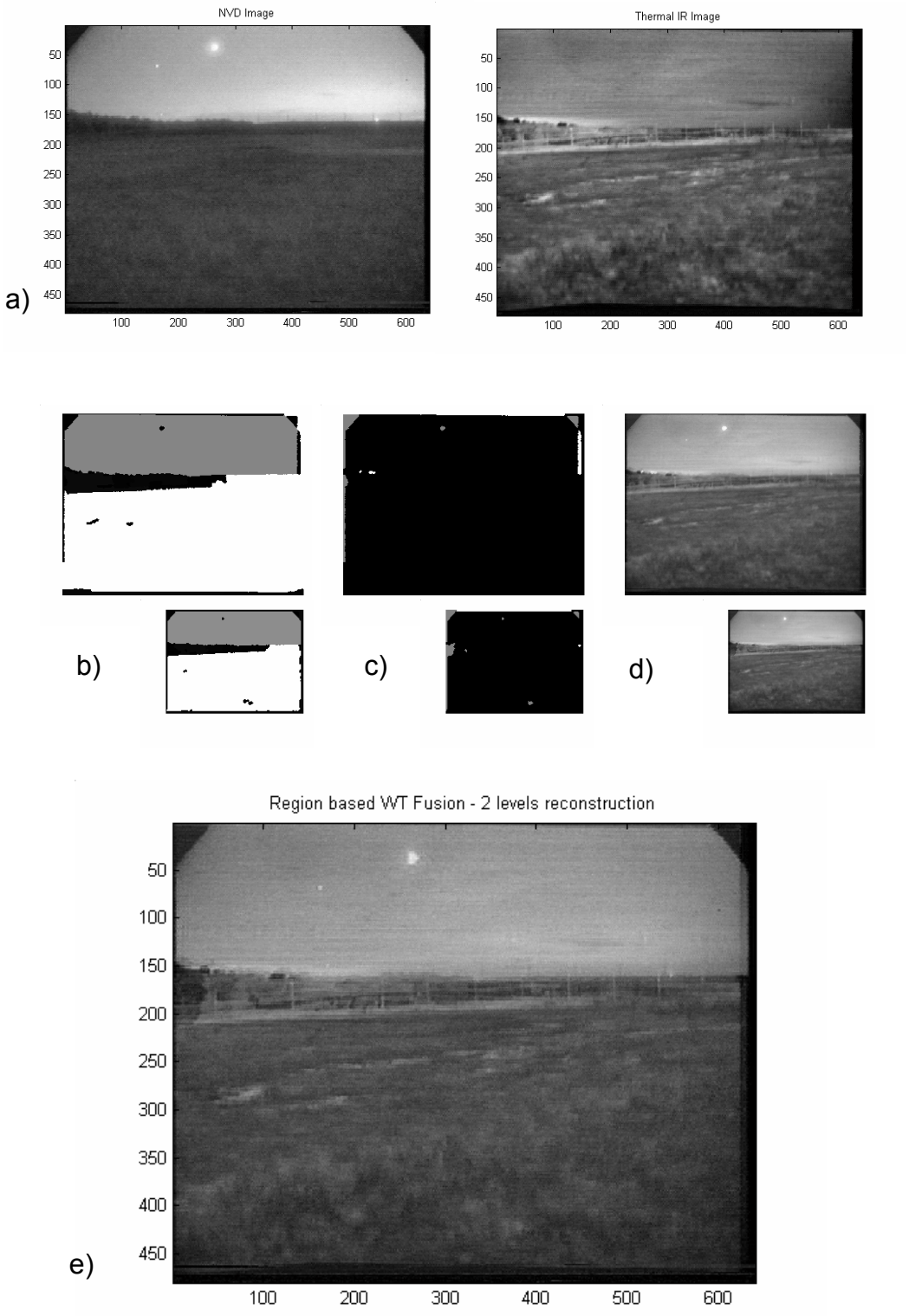


Figure 40. Test 5-2: a) NVD and thermal IR source images; b) Joint region maps achieved using the watershed transform; c) 1st and 2nd level decision maps; d) 1st and 2nd level fused approximate sub-images and e) Reconstructed Fused Image (source images from Naval Research Laboratory).

In summary, the experimental results displayed in both Figure 38 and Figure 40 show that the proposed region-based fusion algorithm retains the most important features from both the Night Vision and thermal IR sensors. For orientation and situation awareness, this is a satisfactory presentation of the datasets and it has improved considerably over the simpler wavelet transform fusion method. Similar to the wavelet-based implementation, further tests and evaluations are needed to determine the optimal settings of the fusion parameters.

V. DISCUSSION AND CONCLUSIONS

This thesis presents a general framework for the multiresolution fusion of NVD and thermal IR imagery. The objective is to exploit the complementary nature of multispectral sensors. The framework encompasses a wavelet-based approach that supports both pixel-level and region-based fusion. The algorithms were tested on different sets of images and the results are evaluated based on a perceptual comparison with the multimodal source images.

In the pixel-level fusion method, variants of the algorithm incorporating different feature selection rules were implemented. By comparing the intensity of the sampled pixels or the activity of a neighborhood (3 by 3 window) around the sampled pixel in the corresponding multiscale wavelet coefficient maps, the salient directional features in the source images can be extracted and selectively combined. The experimental results show that wavelet transform fusion performs better than simple non-multiresolution approaches, e.g., the averaging method and offers significant scene content improvement over single sensor detection. This wavelet-based based approach works well in preserving the key spectral information in the NVD and thermal IR images.

In the wavelet domain, many image processing techniques can easily be performed. Therefore, we propose a region-based fusion scheme, which applies the concept of the watershed transform to the morphological gradient of the decomposed wavelet sub-images. In this approach, the multimodal approximate sub-images are segmented into regions of interest and subsequently used to guide the fusion process. The objective is to increase the degree of subject relevance in the fused image.

Experimental results show that in most cases, the marker-based watershed transform can be used to segment the approximate sub-images into distinct identifiable regions. By considering a region's activity measure in the fusion process, a greater emphasis is placed on the 'regions of interest' representing the salient features in the source images. As a result, the most

important features from the Night Vision and thermal IR sensors are well retained in the fused representation and this scheme leads to a considerable performance improvement over the simpler wavelet transform fusion.

If the segmented regions show similar activity measures, the fused approximate sub-image is obtained by averaging the coefficients of the corresponding source images and the results achieved are comparable to the pixel-level wavelet fusion methods.

Experimental results illustrate the feasibility of the region-based approach for image fusion. The implementation is still at a preliminary stage, and further investigations are proposed to fine tune the approach and vary parameters to improve the fusion performance.

VI. RECOMMENDATIONS FOR FURTHER WORK

Recommended tasks for further research include the following:

- Explore other configurations to determine the optimal settings for both pixel-level and region-based fusion. In this thesis, the Daubechies (db2) wavelets and up to three levels of decomposition are implemented.
- Extend beyond the current fusion scheme (absolute value of wavelet coefficient) by applying more sophisticated criteria, such as a region's size, texture content, and center of mass etc., to further characterize a region's activity level and better reflect a region's relative importance. These parameters can be extracted by examining the magnitude of the wavelet coefficients of each detail sub-band or post-processing the outputs of the watershed transform.
- Explore other fusion rules and methods of multiresolution segmentation, e.g. segmentation based on a generalized pyramid linking [18], hierarchical watershed algorithm from mathematical morphology, etc. The fused results can be compared to determine the most promising approach.
- Examine additional multimodal images, made up of different scenes and targets of interest. This can be done using the newly acquired NVD and thermal cameras acquired in the project. However, images captured with different cameras can no longer be assumed to be registered. Therefore, further study on the registration of the NVD and thermal IR images is necessary.
- Identify suitable applications so that the fusion rules can be automated.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. WAVELET TRANSFORM FUSION RESULTS

The implemented wavelet transform fusion algorithm is tested on additional sets of NVD and thermal IR images (from Naval research Laboratory) having different scene information. The fused results are shown in Figure 41 and Figure 42.

Test A-1

Test Objectives:	To implement and evaluate the performance of wavelet transform fusion on a different pair of NVD and thermal IR images.
Levels of Decomposition:	2 levels
Wavelet family:	Daubechies, db2
Fusion scheme:	Fusion Rule 1 – Selection of the dominant mode

Test A-2

Test Objectives:	To implement and evaluate the performance of wavelet transform fusion on a different pair of NVD and thermal IR images.
Levels of Decomposition:	3 levels
Wavelet family:	Daubechies, db2
Fusion scheme:	Fusion Rule 3 – Weighted average of window-based modes

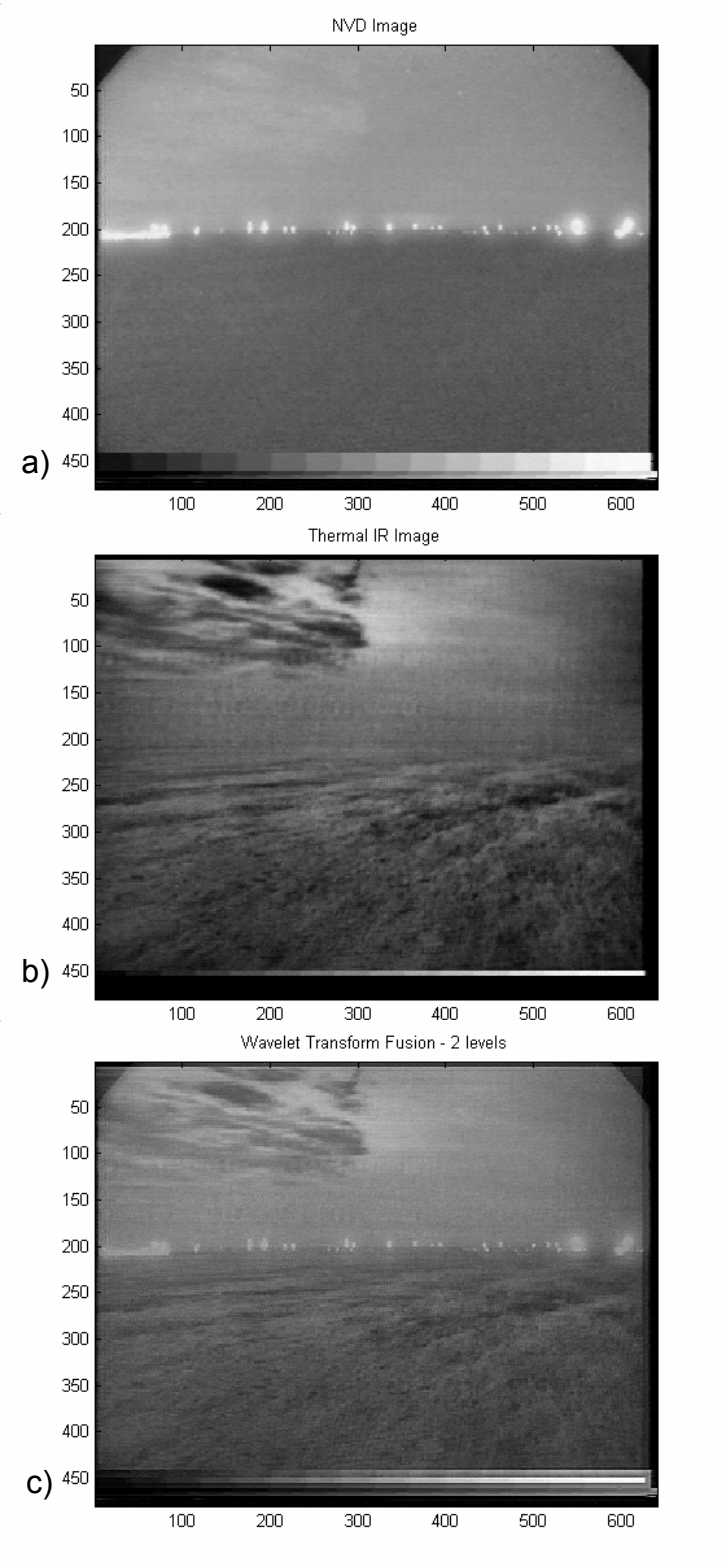


Figure 41. Test A-1 (Wavelet transform fusion results):a) NVD image; b) thermal IR image, and c) wavelet transform fusion with 2 levels of decomposition (source images from Naval Research Laboratory).

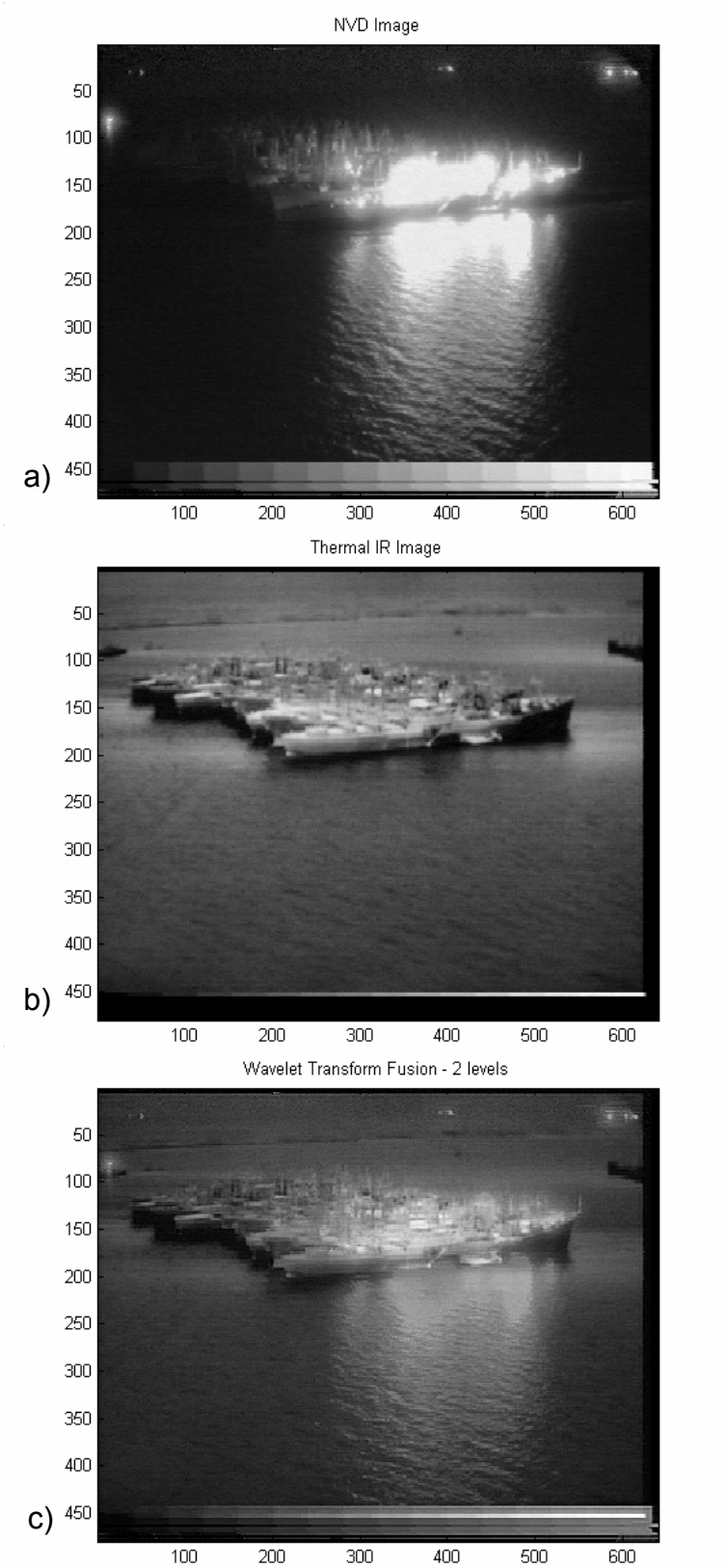


Figure 42. Test A-2 (Wavelet transform fusion results) a) NVD image; b) thermal IR image, and c) Wavelet transform fusion with 2 levels of decomposition (source images from Naval Research Laboratory).

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. REGION FUSION RESULTS

The implemented region-based fusion algorithm is tested on additional sets of NVD and thermal IR images (from Naval research Laboratory) having different scene information. Fused results are shown in Figure 43.

Test B-1

Test Objectives:	To implement and evaluate the performance of the proposed region-based fusion algorithm on a different pair of NVD and thermal IR images.
Levels of Decomposition:	2 levels
Wavelet family:	Daubechies, db2
Fusion scheme:	Region-based fusion rule

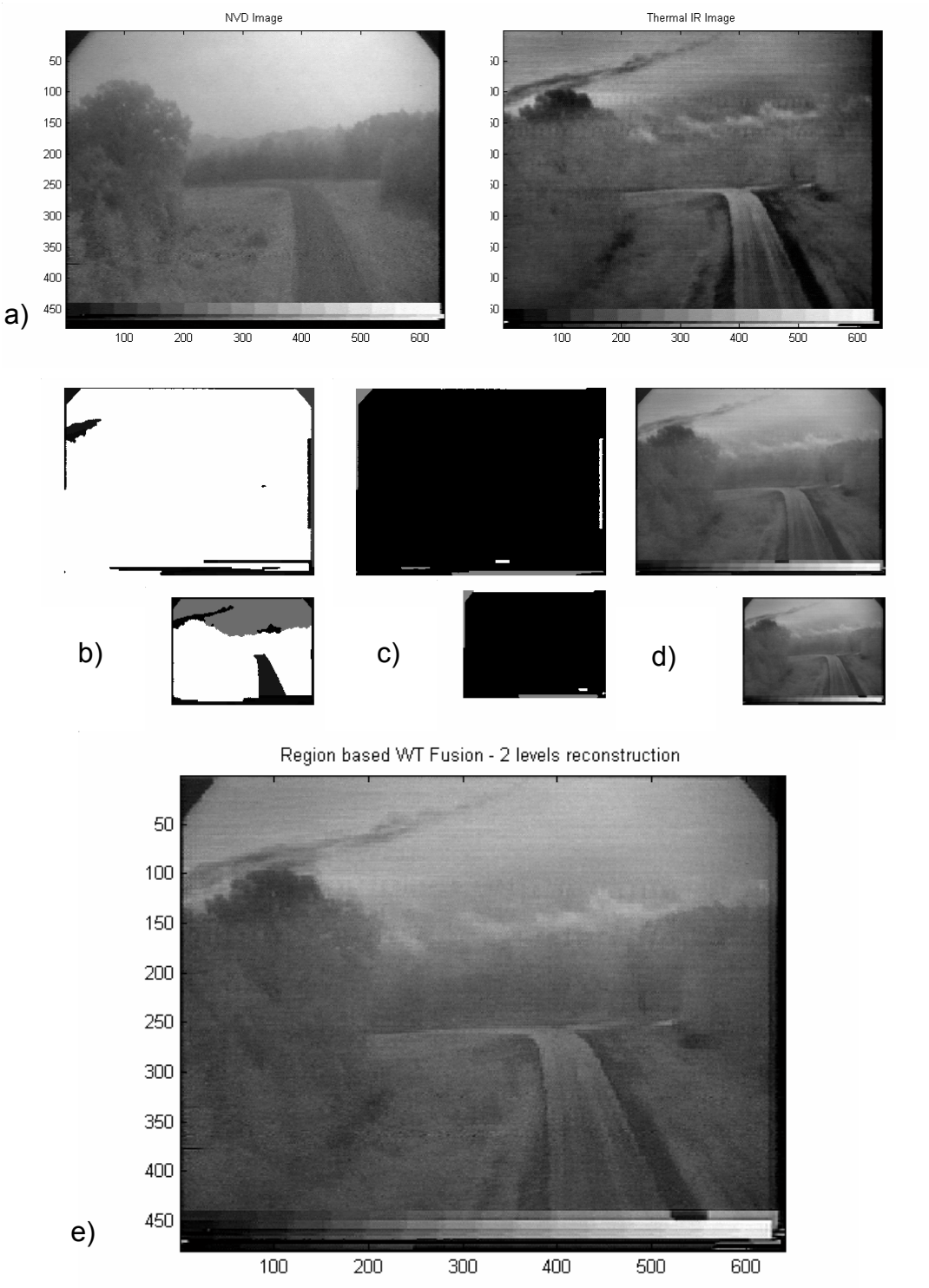


Figure 43. Test B-1 (Region fusion results): a) NVD and thermal IR source images; b) Joint region maps achieved using watershed transform; c) level 1 and 2 decision maps; d) level 1 and 2 fused approximate sub-images and e) reconstructed fused image (source images from Naval Research Laboratory).

LIST OF REFERENCES

1. C.R. Sasso, Major, “Soviet Night Operations in World War II,” Combined Arms Research Library, Command & General Staff College, *Lavenworth Paper* No. 6, December 1982. [http://www-cgsc.army.mil/carl/resources/csi/Sasso/SASSO.asp], October 2002.

2. USAF School of Aerospace Medicine Aeromedical Consultation Service Aerospace Vision Section, “Night Vision Goggles”. [www.brooks.af.mil/web/af/courses/amp/AMP_Online/amp_Eye_lectures/Volume_I/NVG_AMP.doc], October 2004.

3. R.C. Harney, Notes for SI4112 (Combat Systems), “Volume 2. Sensor Elements - Part II Sensor Technologies,” Naval Postgraduate School, January 2004.

4. K-1 Sentry, “Encyclopedia – Thermal Imaging”. [http://www.cctv-information.co.uk/], October 2004.

5. A. Rogalski and K. Chirzanowski, “Infrared devices and techniques,” *Opto-Electronics Review* 10(2), 111-136, 2002.

6. R. McDaniel, D. Scribner, W. Krebs, P. Warren, P. Warren, N. Ockman, J. McCarley, “Image Fusion for Tactical Applications,” *Proceedings of the SPIE - Infrared Technology and Applications XXIV*, Volume 3436, pp. 685-695, 1998.

7. C.W. Therrien, J. Scrofani, and W.K. Krebs, “An adaptive technique for the enhanced fusion of low-light visible with uncooled thermal infrared imagery,” *IEEE International Conference on Imaging Processing*, October 1997.

8. J. Qu and C. Wang, “A Wavelet Package-Based Data Method for Multitemporal Remote Sensing Image Processing,” *22nd Asian Conference on Remote Sensing*, 2001

9. P.J. Burt, “The Pyramid as a Structure For Efficient Computation,” *Multi-resolution Image Processing and Analysis*, pp. 6-35, Berlin: Springer, 1984.

10. P.J. Burt and R.J. Kolczynski, “Enhanced image capture through fusion,” *Fourth International Conference on Computer vision*, IEEE Computer Society, pp. 173-182, 1993.

11. H. Li, B.S. Manjunath, and S.K. Mitra, “Multi-Sensor Image Fusion Using The Wavelet Transform,” *IEEE Computer Society Press*, Los Alamitos, CA, 1994.

12. S.G. Nikolov, P. Hill, D.R. Bull, and C.N. Canagarajah, “Wavelets for Image Fusion,” Wavelets in Signal and Image Analysis: From Theory to Practice edited by Arthur A. Petrosian, October 2001.

13. L.H. David and J. Llinas, *Handbook of Multisensor Data Fusion*, CRC Press, 2001.

14. J.C. Goswami and A.K. Chan, *Fundamentals of Wavelets: Theory, Algorithms and Applications*, Wiley Series in Microwave and Optical Engineering, 1999.

15. Digital Image Processing Documentation for Mathematica – Discrete Wavelet Transform [http://documents.wolfram.com/applications/digitalimage/UsersGuide/8.6.html], September 2004.

16. MathWorks - Wavelets: A New Tool for Signal Analysis [http://www.mathworks.com/access/helpdesk/help/toolbox/wavelet/wavelet.html], September 2004.

17. S.G. Mallat, “A Theory for Multiresolution Signal Decomposition: The Wavelet Representation,” *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Volume 11, No. 7, 1989.

18. G. Piella, “A region-based multiresolution image fusion algorithm,” *Proceedings of the Fifth International Conference on Information Fusion*, 2002.

19. Z. Zhang and R.S. Blum, “A Categorization of Multiscale-Decomposition-Based Image Fusion Schemes with a Performance Study for a Digital Camera Application,” *Proceedings of the IEEE*, Vol 87, August 1999.

20. Z. Fan, S. Fu, R. Li and B. Zuo, “Feature-Level Image Fusion Technique Based on Wavelet Transform,” *Proceedings of SPIE*, Vol 4919, 2002.

21. Z. Zhang and R.S. Blum, “A region-based Image Fusion scheme for concealed weapon detection,” *Proceedings of 31st Annual Conference on Information Sciences and Systems*, March1997.

22. R.C. Gonzalez, R.E. Woods, and S.L. Eddins, *Digital Image Processing using Matlab*, Pearson Prentice Hall, 2004.

23. Z. Shi and R. Shibasaki, “An approach to image segmentation using multiresolution analysis of Wavelets,”, *IEEE International Conference on Systems, Man, and Cybernetics*, 1999.

24. E.R. Dougherty and R.A. Lotufo, *Hands-on Morphological Image Processing*, SPIE Press, 2003.

25. Centre De Morphologie Mathematique, Image Segmentation and Mathematical Morphology, [http://cmm.ensmp.fr/~beucher/wtshed.html], October 2004.

26. SDC Morphological Toolbox for MATLAB, [http://www.mmorph.com/], September 2004.

27. C.R. Jung and J. Scharcanski, “Robust Watershed Segmentation using the Wavelet Transform,” *Proceedings of the WV Brazilian Symposium on Computer Graphics and Image Processing*, 2002.

28. Z. Li, Z. Jing, G. Liu, S. Sun, and H. Leung, “A Region-Based Image Fusion Algorithm using Multiresolution Segmentation,” *The Proceedings of the 2003 IEEE International Conference on Intelligent Transportation Systems*.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Professor Monique P. Fargues
   ECE Department
   Naval Postgraduate School
   Monterey, California

4. Professor Alfred W. Cooper
   Physics Department
   Naval Postgraduate School
   Monterey, California

5. Professor Yeo Tat Soon
   Temasek Defence Systems Institute
   Singapore

6. Mr Chow Khin Choong
   Singapore Technologies Kinetics Ltd
   Singapore